

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2002, 08:23:35 ; Search time 412.25 Seconds
(without alignments)
11544.662 Million cell updates/sec

Title: US-09-898-556A-3

Perfect score: 2772

Sequence: 1 caggcgcttaagctgltg.....ttctaccatctctaccct 2772

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 1000 summaries

Database : N.Geneseq_032802.*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1577	56.9	2614	23	AAST4823 DNA encoding novel
2	571	20.6	678	22	AAAF4909 Human breast cancer
3	525	18.9	653	23	AAAS68256 DNA encoding novel
4	231	8.3	983	22	AAAS26286 Human CDNA encoding novel
5	99	3.6	394	22	AAI83890 Human polynucleotide
6	69	2.5	69	22	AAAI1031 DNA encoding zinc
7	69	2.5	69	22	AAAI1031 DNA encoding zinc
8	49	1.8	821	23	AAAS68257 DNA encoding novel
9	29	1.0	800	22	AAH33832 Human colon cancer

10	29	1.0	2349	23	AAAS90429 DNA encoding novel
11	29	1.0	3582	22	AAIS58664 Human polynucleotide
12	29	1.0	3582	22	AAI60450 Human polynucleotide
13	28	1.0	159	22	AAAL00216 Human reproductive
14	27	1.0	1183	22	AAAS25949 Human CDNA encoding
15	27	1.0	2717	22	AAAS26635 Human genomic DNA
16	27	1.0	3099	22	AAAV64361 Human stem cell z1
17	25	0.9	51	23	ABLO0474 Human silent nonco
18	25	0.9	51	23	ABLO0878 Human amino acid c
19	25	0.9	763	22	AAK92290 Human CDNA 5'-end
20	25	0.9	763	22	AAK93811 Human CDNA clone r
21	25	0.9	814	22	AAI95339 Human neuroblastom
22	25	0.9	871	22	ABAO8823 Human PRO1847 homo
23	25	0.9	2200	22	AAK94575 Human full-length
24	23	0.8	688	22	AAK91560 Human CDNA 5'-end
25	23	0.8	688	22	AAK93129 Human CDNA clone r
26	23	0.8	1512	22	ABA50438 Human breast cell
27	23	0.8	1512	22	ABA68388 Human foetal liver
28	23	0.8	1512	22	ABA35382 Probe #13848 for g
29	23	0.8	1512	22	AAK16761 Human brain expres
30	23	0.8	1512	22	AAK42537 Human bone marrow
31	23	0.8	1512	22	AAI23284 Probe #13217 for g
32	23	0.8	1512	22	AAI48604 Probe #17290 used
33	23	0.8	1512	22	AAI08925 Probe #8916 used t
34	23	0.8	1623	22	AAK94635 Human full-length
35	23	0.8	1965	22	ABA45305 Human breast cell
36	23	0.8	1965	22	ABA55794 Human foetal liver
37	23	0.8	1965	22	ABA25474 Probe #3940 for ge
38	23	0.8	1965	22	AAK04015 Human brain expres
39	23	0.8	1965	22	AAK29501 Human bone marrow
40	23	0.8	1965	22	AAI14072 Probe #4005 for ge
41	23	0.8	1965	22	AAI35452 Probe #4138 used t
42	23	0.8	1965	22	AAI03925 Probe #3916 used t
43	23	0.8	2534	23	AAAS81041 DNA encoding novel
44	22	0.8	439	22	AAAF67397 Novel human polynu
45	22	0.8	484	21	AAAC79234 Human lung tumour-
46	22	0.8	484	21	AAAD2310 Human lung tumour-
47	22	0.8	758	22	AAAS26521 Human CDNA encoding
48	22	0.8	1038	22	ABA21456 Human nervous syst
49	22	0.8	1038	22	AAK84268 Human immune/haema
50	22	0.8	1185	23	AAAS71465 DNA encoding novel
51	22	0.8	1294	23	AAAS64740 DNA encoding novel
52	22	0.8	1661	22	AAAS64739 DNA encoding novel
53	22	0.8	1461	22	AAAS1912 Human polynucleoti
54	22	0.8	1599	23	AAAS71467 DNA encoding novel
55	22	0.8	1812	22	AAK52682 Human polynucleoti
56	22	0.8	2072	22	AAK52896 Human polynucleoti
57	22	0.8	2199	22	AAK51698 Human polynucleoti
58	22	0.8	2259	22	AAAS26098 Human CDNA encoding
59	22	0.8	2382	20	AAAV64360 Human stem cell z1
60	22	0.8	2391	22	AAK94745 Human full-length
61	22	0.8	2472	23	AAAS71466 DNA encoding novel
62	22	0.8	2659	22	AAK52096 Human polynucleoti
63	22	0.8	2659	22	ABA83049 Human transcriptio
64	22	0.8	4029	23	AAAS65913 DNA encoding novel
65	22	0.8	11221	22	AAAS26799 Human genomic DNA
66	21	0.8	264	23	ABLI21853 Drosophila melanog
67	21	0.8	300	21	AAAO0606 Human colon cancer
68	21	0.8	455	23	AAAS86928 DNA encoding novel
69	21	0.8	477	22	AAI92157 Human polynucleoti
70	21	0.8	584	22	AAAI14925 Human breast cancer
71	21	0.8	605	22	AAH07232 Human CDNA clone (
72	21	0.8	693	22	AAH04531 Human CDNA clone (
73	21	0.8	752	21	AAA02263 Human colon cancer
74	21	0.8	781	22	AAI23783 Human breast cancer
75	21	0.8	848	22	AAAS31025 Human diagnostic a
76	21	0.8	874	22	AAH06905 Human CDNA clone (
77	21	0.8	1426	22	AAAS2270 Human polynucleoti
78	21	0.8	1494	24	ABR198977 Human cancer supp
79	21	0.8	1663	20	AAK39662 Renal cancer assoc
80	21	0.8	1705	22	AAK53253 Human polynucleoti
81	21	0.8	1705	22	AAK53254 Human polynucleoti
82	21	0.8	1765	24	AAAI8786 CDNA encoding huma

83	21	0.8	1828	22	AAK52269	Human polynucleoti
84	21	0.8	1964	22	AAK94879	Human full-length
85	21	0.8	2008	22	ABA09199	Human NY-BEN-21 Ag
86	21	0.8	2038	22	AAI60805	Human polynucleoti
87	21	0.8	2038	22	AAI59019	Human polynucleoti
88	21	0.8	2327	23	ABU21852	Drosophila melanog
89	21	0.8	2382	22	AAI66707	A. gossypii AG011
90	21	0.8	2732	22	AAH17926	Human cDNA sequenc
91	21	0.8	2735	22	ABU21824	Drosophila melanog
92	21	0.8	3204	22	AAH17368	Human cDNA sequenc
93	21	0.8	4199	23	ABU22380	Drosophila melanog
94	21	0.8	4829	23	AAK86168	DNA encoding novel
95	21	0.8	6930	22	AAK86280	Human immune/haema
96	21	0.8	10926	22	AAK65370	Human immune/haema
97	20	0.7	127	22	ABA49753	Human breast cell
98	20	0.7	127	22	ABA67665	Human foetal liver
99	20	0.7	127	22	ABA75338	Human foetal liver
100	20	0.7	127	22	ABA34733	Probe #13199 for g
101	20	0.7	127	22	AAK16083	Human brain expres
102	20	0.7	127	22	AAK1825	Human bone marrow
103	20	0.7	127	22	AAI22590	Probe #12523 for g
104	20	0.7	127	22	AAI47880	Probe #16566 used
105	20	0.7	127	22	AAI08268	Probe #8259 used t
106	20	0.7	256	22	AAK31061	Human secreted pro
107	20	0.7	338	21	AAK00814	Human secreted pro
108	20	0.7	345	21	AAK01833	Human reproductive
109	20	0.7	355	21	AAH31025	Human colon cancer
110	20	0.7	370	21	AAK00812	Human secreted pro
111	20	0.7	408	22	AAK31074	Human diagnostic a
112	20	0.7	412	22	ABA09359	Human zn finger pr
113	20	0.7	413	22	AAK67544	Novel human polynu
114	20	0.7	425	22	AAI81543	Human polynucleoti
115	20	0.7	426	22	ABA44606	Human breast cell
116	20	0.7	426	22	ABA55057	Human foetal liver
117	20	0.7	426	22	ABA24812	Probe #3278 for ge
118	20	0.7	426	22	AAK03323	Human brain expres
119	20	0.7	426	22	AAK28774	Human bone marrow
120	20	0.7	426	22	AAI13359	Probe #3282 for ge
121	20	0.7	426	22	AAI134710	Probe #3396 used t
122	20	0.7	426	22	AAI03255	Probe #3246 used t
123	20	0.7	457	22	AAK42486	Human cDNA encodin
124	20	0.7	478	21	AAK00809	Human secreted pro
125	20	0.7	495	21	AAK00810	Human secreted pro
126	20	0.7	495	22	AAK02060	DNA encoding molec
127	20	0.7	518	22	ABA62830	Human foetal liver
128	20	0.7	524	22	AAH97864	Murine 7-transmemb
129	20	0.7	535	22	AAH97863	Murine 7-transmemb
130	20	0.7	570	22	AAH31073	Human diagnostic a
131	20	0.7	585	22	AAH06181	Human cDNA clone (
132	20	0.7	607	23	AAK88222	DNA encoding novel
133	20	0.7	659	22	AAH31068	Human diagnostic a
134	20	0.7	675	21	AAH31140	Human colon cancer
135	20	0.7	798	22	AAK42498	Human cDNA encodin
136	20	0.7	847	22	AAI20037	Human breast cancer
137	20	0.7	888	22	AAH72572	Human cervical can
138	20	0.7	1257	23	AAK86710	DNA encoding novel
139	20	0.7	1257	23	AAK92697	DNA encoding novel
140	20	0.7	1294	22	ABA08437	Human zn finger pr
141	20	0.7	1349	22	AAK03038	Human diagnostic a
142	20	0.7	1705	23	AAK84292	DNA encoding novel
143	20	0.7	1715	23	AAK92934	DNA encoding novel
144	20	0.7	1975	23	AAK87646	DNA encoding novel
145	20	0.7	2097	23	AAK87594	DNA encoding novel
146	20	0.7	2101	21	AAK72427	Human nucleic acid
147	20	0.7	2123	21	AAK92933	DNA encoding novel
148	20	0.7	2181	22	ABA82968	Human transcriptio
149	20	0.7	2228	22	AAK51909	Human polynucleoti
150	20	0.7	2235	23	AAK81692	DNA encoding novel
151	20	0.7	2274	23	AAK56578	DNA encoding novel
152	20	0.7	2313	23	AAK74278	DNA encoding novel
153	20	0.7	2313	23	AAK79107	DNA encoding novel
154	20	0.7	2313	23	AAK92935	DNA encoding novel
155	20	0.7	2345	22	AAH44772	Human DNA mismatch
156	156	20	2386	23	AAK86931	DNA encoding novel
157	157	20	2406	24	ABA95862	Human zinc finger
158	158	20	2478	22	AAH17764	Human cDNA sequenc
159	159	20	2865	22	AAK52893	Human polynucleoti
160	20	0.7	3082	22	AAK53018	Human polynucleoti
161	20	0.7	3342	23	AAK92930	DNA encoding novel
162	20	0.7	3727	22	AAK52034	Human polynucleoti
163	20	0.7	4272	22	AAK58698	Human polynucleoti
164	20	0.7	4420	22	AAK52954	Human polynucleoti
165	20	0.7	4422	22	AAK60484	Human polynucleoti
166	20	0.7	6135	22	ABA20822	Human nervous syst
167	167	20	6135	22	AAK29211	Genomic sequence #
168	168	20	6135	22	AAK85912	Human immune/haema
169	169	20	7470	22	AAK89416	Human digestive sy
170	170	20	7470	22	AAK89417	Human digestive sy
171	171	20	7470	22	AAK89417	Human liver associ
172	172	20	7470	22	AAK31921	Human liver associ
173	173	20	7470	22	AAK31922	Human liver associ
174	174	20	28588	22	AAK89418	Human digestive sy
175	175	20	28588	22	AAK31923	Human liver associ
176	176	19	69	22	AAK513033	DNA encoding zinc
177	177	19	127	22	AAK50423	Human bone marrow
178	178	19	223	21	AAK08609	Human secreted pro
179	179	19	275	22	AAK91335	Human polynucleoti
180	180	19	279	20	AAK214672	Human gene expres
181	181	19	290	21	AAK08608	Human secreted pro
182	182	19	392	21	AAK10140	Human secreted pro
183	183	19	427	23	AAK78109	DNA encoding novel
184	184	19	449	22	AAI92077	Human polynucleoti
185	185	19	451	22	AAI84583	Human polynucleoti
186	186	19	452	22	ABA58066	Human foetal liver
187	187	19	452	22	AAK06138	Human brain expres
188	188	19	452	22	AAK31792	Human bone marrow
189	189	19	452	22	AAI37667	Probe #6353 used t
190	190	19	479	22	AAK06869	Human brain expres
191	191	19	479	22	AAK32591	Human bone marrow
192	192	19	479	22	AAI38410	Probe #7096 used t
193	193	19	538	22	AAK37553	Human bone marrow
194	194	19	581	22	ABA64230	Human foetal liver
195	195	19	581	22	AAK12708	Human brain expres
196	196	19	581	22	AAI19222	Probe #9155 for ge
197	197	19	581	22	AAI44382	Probe #13068 used
198	198	19	597	22	AAI24099	Human breast cancer
199	199	19	599	22	AAK68309	Human lung tumour
200	200	19	633	22	AAK31056	Human diagnostic a
201	201	19	635	22	AAI15253	Human breast cancer
202	202	19	664	22	AAK26518	Human cDNA encodin
203	203	19	692	22	AAK26250	Human cDNA encodin
204	204	19	721	22	AAI95208	Human neuroblastom
205	205	19	753	22	AAH06677	Human cDNA clone (
206	206	19	792	22	AAH06456	Human cDNA clone (
207	207	19	833	22	AAI94504	Human neuroblastom
208	208	19	901	22	AAK31036	Human diagnostic a
209	209	19	957	22	AAK27636	DNA encoding human
210	210	19	976	22	AAK44904	Human contig polyn
211	211	19	991	22	AAK31065	Human diagnostic a
212	212	19	1027	22	AAK31034	Human diagnostic a
213	213	19	1084	23	AAK68741	DNA encoding novel
214	214	19	1106	22	AAK44732	Human full-length
215	215	19	1265	22	ABA51279	Human breast cell
216	216	19	1265	22	ABA69287	Human foetal liver
217	217	19	1265	22	ABA36213	Probe #14679 for g
218	218	19	1265	22	AAK17574	Human brain expres
219	219	19	1265	22	AAK43387	Human bone marrow
220	220	19	1365	22	AAI24167	Probe #14100 for g
221	221	19	1365	22	AAI49454	Probe #18140 used
222	222	19	1365	22	AAI09731	Probe #9722 used t
223	223	19	1390	22	AAH33579	Human colon cancer
224	224	19	1431	22	AAH46940	Human secreted pro
225	225	19	1431	22	AAK52305	Human polynucleoti
226	226	19	1473	22	AAK53289	Human polynucleoti
227	227	19	1517	22	AAI58016	Human polynucleoti
228	228	19	1561	24	AAH45687	Human zinc finger
		19			AAK24005	Human inflammation

375	18	0.6	577	22	AAS26358	Human cDNA encodin	448	18	0.6	1785	22	AAH75047	Nucleotide sequenc
376	18	0.6	584	23	AAS68996	DNA encoding novel	449	18	0.6	1785	22	AAH76370	Human PMSR3 cDNA.
377	18	0.6	594	22	ABA61905	Human foetal liver	450	18	0.6	1810	21	AAH44977	Arabidopsis thalia
378	18	0.6	594	22	ABA29442	Probe #7908 for ge	451	18	0.6	1817	22	AAK94651	Human full-length
379	18	0.6	594	22	AAK10216	Human brain expres	452	18	0.6	1818	23	ABL22397	Drosophila melanog
380	18	0.6	594	22	AAK36114	Human bone marrow	453	18	0.6	1865	23	AAS68335	DNA encoding novel
381	18	0.6	594	22	AAI17236	Probe #7169 for ge	454	18	0.6	1865	23	AAS69005	DNA encoding novel
382	18	0.6	594	22	AAI41830	Probe #10516 used	455	18	0.6	1883	21	AAS672418	Human nucleic acid
383	18	0.6	614	23	AAS25989	Human cDNA encodin	456	18	0.6	1999	21	AAH72418	Human cDNA encodin
384	18	0.6	614	23	AAS58429	DNA encoding novel	457	18	0.6	2000	20	AAZ20098	Rice serine palmit
385	18	0.6	620	23	AAS70669	DNA encoding novel	458	18	0.6	2038	22	AAS22799	Human cDNA encodin
386	18	0.6	638	22	AAS26425	Human cDNA encodin	459	18	0.6	2133	9	AAV01890	Human OFK18 gene.
387	18	0.6	678	22	ABA08751	Human secreted pro	460	18	0.6	2156	18	AAV01890	Sequence of Haell I
388	18	0.6	738	20	AAK98852	Human secreted pro	461	18	0.6	2204	22	AAS26025	Human cDNA encodin
389	18	0.6	738	22	AAK83140	Human validated ca	462	18	0.6	2361	22	AAH14258	Human cDNA sequenc
390	18	0.6	738	22	AAK83141	Human immune/haema	463	18	0.6	2369	24	ABA01156	Human zinc finger
391	18	0.6	740	13	AAO27236	DNA encoding huma	464	18	0.6	2412	22	AAS22563	Human cDNA encodin
392	18	0.6	748	20	AAK61450	DNA encoding a hum	465	18	0.6	2427	22	AAS200460	Plasmid pLO1223 u
393	18	0.6	781	21	AAK77651	Human cancer assoc	466	18	0.6	2647	21	AAA30827	Expression cassett
394	18	0.6	795	22	AAH75048	Nucleotide sequenc	467	18	0.6	2720	22	AAH72833	Secreted protein g
395	18	0.6	795	23	AAH6371	Human PMS19 cDNA.	468	18	0.6	2852	21	AAK68501	Human immune/haema
396	18	0.6	795	23	AAS64690	DNA encoding novel	469	18	0.6	2880	21	AAA23438	Human cDNA encodin
397	18	0.6	799	22	ABA49151	Human breast cell	470	18	0.6	2890	22	AAK93782	Human cDNA encodin
398	18	0.6	799	22	ABA67064	Human foetal liver	471	18	0.6	2890	22	AAK93782	Human cDNA encodin
399	18	0.6	799	22	ABA34157	Probe #12623 for g	472	18	0.6	2913	22	AAK51782	Human poly nucleot
400	18	0.6	799	22	AAK15508	Human brain expres	473	18	0.6	2924	22	AAH72831	Secreted protein g
401	18	0.6	799	22	AAK41238	Human bone marrow	474	18	0.6	2926	22	AAH72805	Secreted protein g
402	18	0.6	799	22	AAI21990	Probe #11923 for g	475	18	0.6	2926	22	AAH72830	Secreted protein g
403	18	0.6	799	22	AAI47280	Probe #15966 used	476	18	0.6	2950	21	AAK59988	Human secreted pro
404	18	0.6	799	22	AAI07685	Probe #7676 used t	477	18	0.6	2950	21	AAK26427	Human secreted pro
*405	18	0.6	898	22	AAK91639	Human cDNA 5'-end	478	18	0.6	2951	22	AAH72834	Secreted protein g
406	18	0.6	898	22	AAK93578	Human cDNA clone r	479	18	0.6	2951	22	AAH72832	Secreted protein g
407	18	0.6	908	21	AAC76458	Human OREX ORF2013	480	18	0.6	3033	22	AAH45127	Human TANGO 273 cD

521	17	0.6	53	21	AA29670	Plasmid construct	c 594	17	0.6	596	21	AA12728	Aspergillus oryzae
522	17	0.6	69	22	AA513071	DNA encoding zinc	595	17	0.6	596	22	AA526392	Human cDNA encodin
523	17	0.6	194	22	AB475667	Human foetal liver	596	17	0.6	597	22	AA526251	PERV-MSNA clone Bf
524	17	0.6	194	22	AAK24279	Human brain expres	597	17	0.6	598	20	AA97889	Human secreted pro
525	17	0.6	194	22	AAK50304	Human bone marrow	598	17	0.6	624	22	AA539408	Human secreted pro
526	17	0.6	194	22	AA156271	Probe #24957 used	599	17	0.6	624	22	AAK88256	CDNA encoding nove
527	17	0.6	233	21	AAAC1623	Human secreted pro	600	17	0.6	632	22	AAH97876	Human digestive sy
528	17	0.6	257	21	AA42043	Human secreted exp	601	17	0.6	644	22	AA545029	Murine 7-transemb
529	17	0.6	289	19	AAV60049	Nucleic acid 645.	602	17	0.6	657	20	AA97842	CNA encoding nove
530	17	0.6	297	21	AAA44267	Human secreted pro	603	17	0.6	657	20	AAH44439	Human secreted pro
531	17	0.6	314	21	AAAC00598	Human secreted exp	604	17	0.6	673	22	AAH44439	Human alkylated-DN
532	17	0.6	318	20	AA40476	Human secreted pro	605	17	0.6	676	22	AAE27650	Human musculoskele
533	17	0.6	324	22	AA544566	Human cDNA encodin	606	17	0.6	694	22	AA526254	DNA encoding human
534	17	0.6	343	22	AA165415	Nucleotide sequenc	607	17	0.6	701	20	AA215844	Human cDNA encodin
535	17	0.6	347	20	AA51882	Human secreted pro	608	17	0.6	712	21	AA215844	Human gene express
536	17	0.6	355	22	AA537836	Novel human diagno	609	17	0.6	714	22	AAK57598	Arabidopsis thalia
537	17	0.6	373	23	AA567212	DNA encoding novel	610	17	0.6	714	22	AAK57598	Human immune/haema
538	17	0.6	401	22	ABAI2150	Human nervous syst	611	17	0.6	715	22	AAK70697	Human immune/haema
539	17	0.6	401	22	AAH99901	Human protein enco	612	17	0.6	732	20	AAK25548	Rice abscisic acid
540	17	0.6	413	22	AA192471	Human polynucleoti	613	17	0.6	733	21	AAK76296	Human OREF ORF1851
541	17	0.6	417	20	AA591233	T. gondii immunoge	614	17	0.6	739	22	AAH34145	Human colon cancer
542	17	0.6	417	20	AA542556	T. gondii DNA enco	615	17	0.6	740	22	AAH05067	Human colon cancer
543	17	0.6	441	22	ABA08283	Human Zn finger pr	616	17	0.6	742	22	AAH07538	Human cDNA clone (
544	17	0.6	448	22	ABA57646	Human foetal liver	617	17	0.6	759	22	AAH07538	Human cDNA clone (
545	17	0.6	448	22	ABA27066	Probe #5532 for ge	618	17	0.6	803	22	AAH03899	Human cDNA clone (
546	17	0.6	448	22	AAK05700	Human brain expres	619	17	0.6	836	22	AA526031	Human neuroblastom
547	17	0.6	448	22	AAK31317	Human bone marrow	620	17	0.6	841	22	AAH05482	Human cDNA clone (
548	17	0.6	448	22	AA115632	Probe #5565 for ge	621	17	0.6	847	22	AAH03453	Human cDNA clone (
549	17	0.6	448	22	AA137213	Probe #5899 used t	622	17	0.6	860	22	AA531058	Human diagnostic a
550	17	0.6	470	22	ABA26040	Probe #4506 for ge	623	17	0.6	877	22	AAH06903	Human cDNA clone (
551	17	0.6	470	22	AAK30084	Human bone marrow	624	17	0.6	885	20	AA578655	Human cDNA clone (
552	17	0.6	470	22	AAAD07626	Human secreted pro	625	17	0.6	889	23	AA578655	DNA encoding novel
553	17	0.6	473	22	AAK31276	Human bone marrow	626	17	0.6	905	22	AA526462	Rat U3 gene trap d
554	17	0.6	473	22	AAK63970	Human immune/haema	627	17	0.6	906	22	ABA51548	Human cDNA encodin
555	17	0.6	478	22	ABA52068	Human foetal liver	628	17	0.6	906	22	ABA59610	Human breast cell
556	17	0.6	478	22	AAK21882	Probe #348 for gen	629	17	0.6	906	22	AA124444	Human foetal liver
557	17	0.6	478	22	AAK00350	Human brain expres	630	17	0.6	914	19	AAV29574	Probe #14377 for g
558	17	0.6	478	22	AAK25794	Human bone marrow	631	17	0.6	930	21	AAK48843	Phaffia rhodozyma
559	17	0.6	478	22	AA110423	Probe #356 for gen	632	17	0.6	938	21	AAK48843	Arabidopsis thalia
560	17	0.6	478	22	AA131675	Probe #361 used to	633	17	0.6	953	22	AAK75394	Human OREF ORF949
561	17	0.6	478	22	AA100358	Probe #349 used to	634	17	0.6	972	22	AAH00777	Absetia corymbifer
562	17	0.6	480	22	AAK29994	Human cDNA 3'-end	635	17	0.6	1001	22	AA525928	Human cDNA encodin
563	17	0.6	505	22	ABA61306	Human foetal liver	636	17	0.6	1001	22	ABA49927	Human breast cell
564	17	0.6	505	22	AAK09603	Human brain expres	637	17	0.6	1001	22	ABA67846	Human foetal liver
565	17	0.6	505	22	AAK35496	Human bone marrow	638	17	0.6	1019	22	ABA34905	Probe #13371 for g
566	17	0.6	505	22	AA141210	Probe #9896 used t	639	17	0.6	1029	23	AAK44366	Human bone marrow
567	17	0.6	512	22	AAH09437	Human cDNA clone (640	17	0.6	1037	22	AAH52555	Drosophila melanog
568	17	0.6	516	22	AAK60837	Human immune/haema	641	17	0.6	1047	22	AAH52555	S. epidermidis ope
569	17	0.6	518	22	AAH97872	Murine 7-transemb	642	17	0.6	1065	21	AA503034	Human diagnostic a
570	17	0.6	526	21	AAK39190	Zea mays DNA fragm	643	17	0.6	1097	22	ABA08482	Arabidopsis thalia
571	17	0.6	532	21	AAZ80564	Human colon cancer	644	17	0.6	1103	21	AAK69445	Human Ca channel s
572	17	0.6	541	21	AAK69771	Human breast tumou	645	17	0.6	1123	22	AA510875	Human cDNA encodin
573	17	0.6	544	22	AAK92455	Human cDNA 3'-end	646	17	0.6	1124	22	ABA16976	Human cDNA encodin
574	17	0.6	561	21	AAK69744	Human ovarian carc	647	17	0.6	1125	22	ABA16977	Human nervous syst
575	17	0.6	563	21	AAAI6428	Human colon cancer	648	17	0.6	1129	22	AA545217	Human cDNA encodin
576	17	0.6	566	22	ABA63219	Human foetal liver	649	17	0.6	1140	22	AAH68001	C. glutamicum codin
577	17	0.6	566	22	AAK11684	Human brain expres	650	17	0.6	1152	22	AAH53316	S. epidermidis ope
578	17	0.6	566	22	AAK37427	Human bone marrow	651	17	0.6	1155	21	AAK49626	Arabidopsis thalia
579	17	0.6	566	22	AA143285	Probe #11972 used	652	17	0.6	1173	22	ABA07013	Human pancreatic c
580	17	0.6	567	22	ABA62359	Human foetal liver	653	17	0.6	1173	22	AA532373	Human cDNA encodin
581	17	0.6	567	22	ABA28699	Probe #8165 for ge	654	17	0.6	1183	21	AAK43780	Human cDNA encodin
582	17	0.6	567	22	AAK10694	Human brain expres	655	17	0.6	1202	21	AAK54424	Arabidopsis thalia
583	17	0.6	567	22	AAK36581	Human bone marrow	656	17	0.6	1219	22	AAE27657	DNA encoding human
584	17	0.6	567	22	AA117433	Probe #7366 for ge	657	17	0.6	1241	23	AA573410	DNA encoding novel
585	17	0.6	567	22	AA142339	Probe #11025 used	658	17	0.6	1294	22	AAK85176	Human immune/haema
586	17	0.6	573	22	AA135295	Human musculoskele	659	17	0.6	1297	22	AAH00481	Human secreted pro
587	17	0.6	594	22	AAH10991	Human cDNA clone (660	17	0.6	1299	20	AAK19492	Human alkylated-DN
588	17	0.6	595	22	ABA63831	Human foetal liver	661	17	0.6	1300	22	AA510869	Human cDNA encodin
589	17	0.6	595	22	ABA31012	Probe #9478 for ge	662	17	0.6	1318	18	AA51254	Human AD4 gene gen
590	17	0.6	595	22	AAK12342	Human brain expres	663	17	0.6	1320	20	AA578761	DNA encoding novel
591	17	0.6	595	22	AAK38061	Human bone marrow	664	17	0.6	1329	20	AAZ30546	Micrococcus luteus
592	17	0.6	595	22	AA118833	Probe #8766 for ge	665	17	0.6	1380	21	AAZ70124	Plasmodium falcipar
593	17	0.6	595	22	AA143956	Probe #12642 used	666	17	0.6	1383	22	AAAL07015	Human reproductive

813	17	0.6	2506	22	AAH89919	Human bone marrow
814	17	0.6	2544	23	AAS82160	DNA encoding novel
815	17	0.6	2553	24	ABA05651	Human zinc finger
816	17	0.6	2560	22	AAK53091	Human polynucleoti
817	17	0.6	2561	22	AAH17818	Human cDNA sequenc
818	17	0.6	2565	21	AAZ96805	Nuclear transpor
819	17	0.6	2572	22	AAK53098	Human polynucleoti
820	17	0.6	2572	23	AAK51118	DNA encoding novel
821	17	0.6	2604	22	AAS92936	DNA encoding novel
822	17	0.6	2624	23	AAK52114	Human polynucleoti
823	17	0.6	2625	22	AAK44597	Human full-length
824	17	0.6	2644	22	AAK44769	Human conlig polyn
825	17	0.6	2680	19	AAV64579	MYC-binding zinc-f
826	17	0.6	2694	21	AAZ96806	Nuclear transport
827	17	0.6	2718	23	AAS69395	DNA encoding novel
828	17	0.6	2741	21	AAZ52436	HTPM clone 1484257
829	17	0.6	2761	22	AAI60214	Human polynucleoti
830	17	0.6	2769	22	AAI58428	Human polynucleoti
831	17	0.6	2864	22	AAK52107	Human polynucleoti
832	17	0.6	2920	23	AAS85209	DNA encoding novel
833	17	0.6	2940	24	AB199447	Mouse ischaemic co
834	17	0.6	2959	22	AAK51505	Human polynucleoti
835	17	0.6	2959	22	AAK52489	Human polynucleoti
836	17	0.6	2960	23	ABLI5944	Drosophila melanog
837	17	0.6	2982	21	AAZ7388	AAAT7388
838	17	0.6	3006	21	AAZ37101	CDNA encoding a hu
839	17	0.6	3027	18	AAAT75305	Nucleotide sequenc
840	17	0.6	3041	12	AAQ11852	Glutamate receptor
841	17	0.6	3044	18	AAAT64551	Tomato S-ribonucle
842	17	0.6	3105	24	ABA03746	Human zinc finger
843	17	0.6	3150	21	AACT6591	Human ORF ORF2546
844	17	0.6	3154	22	ABA07338	Human pancreatic c
845	17	0.6	3154	22	AAK32765	Human genomic DNA
846	17	0.6	3166	22	AAS22919	DNA encoding novel
847	17	0.6	3175	22	AAI59723	Human polynucleoti
848	17	0.6	3189	23	AAS85210	DNA encoding novel
849	17	0.6	3197	22	AAS26673	Human genomic DNA
850	17	0.6	3202	22	AAH54022	S. epidermidis gen
851	17	0.6	3213	22	AAI57937	Human polynucleoti
852	17	0.6	3244	23	ABLO8432	Drosophila melanog
853	17	0.6	3264	23	ABLO6061	Drosophila melanog
854	17	0.6	3320	18	AAV09698	Porcine retrovirus
855	17	0.6	3331	23	AAK56067	DNA encoding novel
856	17	0.6	3369	22	AAI26673	Human breast cance
857	17	0.6	3387	22	AAH76384	Human T1LC polypep
858	17	0.6	3434	23	AAH73872	DNA encoding novel
859	17	0.6	3481	22	AAH72651	Human cervical can
860	17	0.6	3497	22	AAH55010	S. epidermidis gen
861	17	0.6	3545	23	ABLO7616	Drosophila melanog
862	17	0.6	3600	20	AAK25180	HIV-1 group O isol
863	17	0.6	3639	23	AAK72011	DNA encoding novel
864	17	0.6	3651	22	AAI36943	Human musculoskele
865	17	0.6	3703	22	AAH44849	Human cDNA encodin
866	17	0.6	3726	23	AAK56684	DNA encoding novel
867	17	0.6	3840	24	AB19881	Mouse ischaemic co
868	17	0.6	3871	23	ABLI0528	Drosophila melanog
869	17	0.6	3900	22	AAH76383	Human T1LC polypep
870	17	0.6	3906	22	AAH73707	Human zinc finger
871	17	0.6	4008	22	ABA15890	Human nervous syst
872	17	0.6	4012	22	ABA07314	Human pancreatic c
873	17	0.6	4012	22	ABA07315	Human pancreatic c
874	17	0.6	4012	22	AAK90473	Human digestive sy
875	17	0.6	4012	22	AAK90474	Human digestive sy
876	17	0.6	4061	21	AAA49923	Human calcium chan
877	17	0.6	4086	23	AAK585765	DNA encoding novel
878	17	0.6	4156	22	AAK83412	Human immune/haema
879	17	0.6	4156	23	ABLI18176	Drosophila melanog
880	17	0.6	4182	22	AAH26524	Human prolon/Oligo
881	17	0.6	4358	22	AAH14605	Human cDNA sequenc
882	17	0.6	4359	22	AAI05832	Human reproductive
883	17	0.6	4402	22	AAK67020	PEVY env protein c
884	17	0.6	4416	21	AAA09335	Human cancer assoc
885	17	0.6	4416	21	AAA09345	Human cancer assoc
886	17	0.6	4527	22	AAI58092	Human polynucleoti
887	17	0.6	4590	21	AACT5767	Human ORF ORF1322
888	17	0.6	4768	21	AAK69134	Human ABC1 gene ex
889	17	0.6	4918	22	AAK67022	PEVY env protein c
890	17	0.6	5167	22	AAK32675	Human genomic DNA
891	17	0.6	5167	22	AAK76525	Human immune/haema
892	17	0.6	5223	23	ABLI3306	Drosophila melanog
893	17	0.6	5440	22	AAI05240	Human reproductive
894	17	0.6	5440	22	AAI05242	Human reproductive
895	17	0.6	5706	22	AAK76497	Human immune/haema
896	17	0.6	5770	23	AAK84301	DNA encoding novel
897	17	0.6	5980	24	ABLI33188	Human immune syste
898	17	0.6	6070	22	AAK81491	Human immune syste
899	17	0.6	6076	22	AAK67021	PEVY env protein c
900	17	0.6	6076	22	AAK67032	PEVY env protein c
901	17	0.6	6173	22	AAK30504	DNA encoding novel
902	17	0.6	6173	22	AAI06284	Human reproductive
903	17	0.6	6173	22	AAK69343	Human immune/haema
904	17	0.6	6174	22	AAK30503	DNA encoding novel
905	17	0.6	6174	22	AAI06283	Human reproductive
906	17	0.6	6174	22	AAK69342	Human immune/haema
907	17	0.6	6393	22	AAK26744	Human genomic DNA
908	17	0.6	6426	22	AAK26754	Human genomic DNA
909	17	0.6	6427	22	AAK26755	Human genomic DNA
910	17	0.6	6689	23	ABLI19618	Drosophila melanog
911	17	0.6	6815	22	AAK26745	Human genomic DNA
912	17	0.6	6815	22	AAK68521	Human immune/haema
913	17	0.6	6916	23	ABLI19616	Drosophila melanog
914	17	0.6	6953	17	AAI17116	Rhodopsin gene. H
915	17	0.6	7120	23	ABLO4480	Drosophila melanog
916	17	0.6	7149	23	ABLI19614	Drosophila melanog
917	17	0.6	7207	21	AAK21316	Human Low adenosin
918	17	0.6	7207	21	AAK35194	Human adenosine re
919	17	0.6	7297	22	AAK89535	Human digestive sy
920	17	0.6	7333	22	AAK77726	Defective retrovir
921	17	0.6	7362	22	AAK67019	PEVY env protein c
922	17	0.6	7393	18	AAK74883	Porcine retrovirus
923	17	0.6	7445	22	ABA07339	Human pancreatic c
924	17	0.6	7445	22	AAK32766	Human genomic DNA
925	17	0.6	7562	23	ABLI18203	Drosophila melanog
926	17	0.6	7873	22	AAK67023	PEVY env protein c
927	17	0.6	8196	18	AAV09699	Porcine retrovirus
928	17	0.6	8209	18	AAV09700	Porcine retrovirus
929	17	0.6	8353	22	AAS26697	Human genomic DNA
930	17	0.6	8372	22	AAK46354	Tumour suppressor
931	17	0.6	8420	22	AAK89536	Human digestive sy
932	17	0.6	8446	24	ABLI33670	Human immune syste
933	17	0.6	8456	24	ABLI33877	Human immune syste
934	17	0.6	9078	23	ABLI13626	Drosophila melanog
935	17	0.6	9183	23	ABLI06060	Drosophila melanog
936	17	0.6	9437	23	ABLI10030	Drosophila melanog
937	17	0.6	9519	22	AAI37169	Human musculoskele
938	17	0.6	9618	22	AAK26698	Human genomic DNA
939	17	0.6	9666	24	ABLI33084	Human immune syste
940	17	0.6	9666	24	ABLI34530	Human metastasis a
941	17	0.6	10205	23	ABLI18202	Drosophila melanog
942	17	0.6	10726	23	ABLO6616	Drosophila melanog
943	17	0.6	10782	22	AAK75423	Human immune/haema
944	17	0.6	11089	22	AAK79653	Human immune/haema
945	17	0.6	11340	22	AAK75424	Human immune/haema
946	17	0.6	11754	22	AAK50433	Human ABC1 gene, p
947	17	0.6	12221	22	AAK541687	Genomic sequence #
948	17	0.6	12221	22	AAK81809	Human immune/haema
949	17	0.6	12221	22	AAK84385	Human immune/haema
950	17	0.6	12584	22	AAK61399	Human immune/haema
951	17	0.6	12586	22	AAK67076	Human gene regulat
952	17	0.6	12806	22	ABLI24266	Human immune/haema
953	17	0.6	13160	22	ABLI5804	Drosophila melanog
954	17	0.6	13189	22	ABLI5805	Human nervous syst
955	17	0.6	13418	22	AAK86473	Human nervous syst
956	17	0.6	13526	22	AAK91515	Human immune/haema
957	17	0.6	13526	22	AAK32169	Human digestive sy
958	17	0.6	14332	22	AAK41688	Human liver associ
						Genomic sequence #

959	17	0.6	14332	22	AAK81810	Human immune/haema
960	17	0.6	14332	22	AAK84386	Human immune/haema
961	17	0.6	14337	22	AAK91514	Human digestive sy
962	17	0.6	14337	22	AAK32168	Human liver associ
963	17	0.6	15254	22	AAK83135	Human immune/haema
964	17	0.6	15254	22	AAK83137	Human immune/haema
965	17	0.6	15255	22	AAK83136	Human immune/haema
966	17	0.6	15256	22	AAK83134	Human immune/haema
967	17	0.6	15418	21	AAK63785	Nucleotide sequenc
968	17	0.6	15630	21	AAK21317	Human low adenosin
969	17	0.6	15630	21	AAK35195	Human adenosine re
970	17	0.6	16271	22	ABA19209	Human nervous syst
971	17	0.6	16271	22	ABA19210	Human nervous syst
972	17	0.6	18663	22	AAK57589	Human immune/haema
973	17	0.6	18996	22	AAK57112	DNA encoding Droso
974	17	0.6	18996	23	ABL11802	Drosophila melanog
975	17	0.6	20811	23	ABL15604	Drosophila melanog
976	17	0.6	21509	23	ABL08688	Human reproductive
977	17	0.6	22635	22	AAH07203	Human reproductive
978	17	0.6	26040	22	AAH27887	Nucleotide sequenc
979	17	0.6	26555	22	AAK68372	Human immune/haema
980	17	0.6	26555	22	AAK68605	Human immune/haema
981	17	0.6	26555	22	AAK62833	Human genomic DNA
982	17	0.6	32194	22	AAH04340	Human reproductive
983	17	0.6	33923	22	AAK67071	Human immune/haema
984	17	0.6	43804	20	AAK76375	Chicken embryo let
985	17	0.6	43804	20	AAK26690	Complete genome se
986	17	0.6	44018	22	AAK82392	Avian adenovirus C
987	17	0.6	49136	21	AAK27475	NIDDM1 region incl
988	17	0.6	56632	22	AAK65581	Human immune/haema
989	17	0.6	72928	20	AAK18355	Human ASTRIJ 5' ge
990	17	0.6	72928	21	AAK80253	Human ASTRIJ 5' ge
991	17	0.6	80578	22	AAH44800	Human GPCR protein
992	17	0.6	81001	22	AAK30035	Human apolipoprote
993	17	0.6	90050	21	AAK91925	Wild type (C57BL/6
994	17	0.6	160271	22	AAK85750	Bipolar affective
995	17	0.6	160271	22	AAK85756	Human chromosome 1
996	17	0.6	160271	22	AAK504858	Human chromosome 1
997	17	0.6	160271	22	AAK504864	Human chromosome 1
998	17	0.6	160271	22	AAK506667	Human chromosome 1
999	17	0.6	160271	22	AAH40997	160kb fragment of
1000	17	0.6	160271	22	AAH23764	Human chromosome 1

ALIGNMENTS

RESULT 1

ID AAS74823 standard; cDNA; 2614 BP.

AC AAS74823;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #10627.

KW Human: chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSB-) HYSBO INC.

XX

PI Drmanac RT, Liu C, Tang YT:
 XX WPI, 2001-639362/73.
 DR P-SDB; ABG10636.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS
 XX
 PS Claim 1: SEQ ID No 10627; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 2614 BP; 745 A; 611 C; 660 G; 598 T; 0 other;

Query Match 56.9%; Score 1577; DB 23; Length 2614;
 Best local Similarity 99.9%; Pred. No. 0;
 Matches 1817; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY	427	cagaatcgaagccagaatccaacttagtcctcctcgtccctcgtatttctccagtcagc	486
DB	229		288
QY	487	aagcttcgaagccagaatccaacttagtcctcctcgtccctcgtatttctccagtcagc	546
DB	289		348
QY	547	cagg-aaatcctcctcgaagccagaatccaacttagtcctcctcgtatttctccagtcagc	605
DB	349		408
QY	606	ccattctgctttagtgcgaagccagaatccaacttagtcctcctcgtatttctccagtcagc	665
DB	409		468
QY	666	tttggagagtaagcaaaatgacacttcaagccacttccagccacttgaagaaca	725
DB	469		528
QY	726	cagccagcagtcctcgaagccagaatccaacttagtcctcctcgtatttctccagtcagc	785
DB	529		588
QY	786	agggcagatctagaagcaaaacagcaagatcgatgcttgaagctccagatttga	845
DB	589		648
QY	846	gaatcaaatatgaagatttggccaggtttataaagagacaaacctcttagccctc	905
DB	649		708

OY	906	cagaagacacaa	actvgvgaaac	acttatacgt	taactgtagtvgvgagaaag	ctttggc	965	
Dp	709	cagaagacacaa	actvgvgaaac	acttatacgt	taactgtagtvgvgagaaag	ctttggc	768	
OY	966	agatgacagtc	ccatcaca	aaaccccaag	gacacactcggvggaaag	cttattgtgc	1025	
Dp	769	agatgvcag	lucctcaca	aaaccccaag	gacacactcggvggaaag	cttattgtgc	828	
OY	1026	agggaaatg	vgvgcgaag	ctttagtvg	aaagctttagtacaacatcagaagaaac		1085	
Dp	829	agggaaatg	vgvgcgaag	ctttagtvg	aaagctttagtacaacatcagaagaaac		888	
OY	1086	tcaagvgagaa	accttattg	tvgaaagat	tgtvgacgaag	cttacttvgaaatcgaaac	1145	
Dp	889	tcaagvgagaa	accttattg	tvgaaagat	tgtvgacgaag	cttacttvgaaatcgaaac	948	
OY	1146	ctctttaca	atcgaac	cccttaaggg	ctcaagcctttagtvtgaaagaaatg	vgvg	1205	
Dp	949	ctctttaca	atcgaac	cccttaaggg	ctcaagcctttagtvtgaaagaaatg	vgvg	1008	
OY	1206	cagaagcttag	cgcgaag	gtcaaaac	cccttatcaacacgaagvgcgaca	actvtgvgaaag	1265	
Dp	1009	cagaagcttag	cgcgaag	gtcaaaac	cccttatcaacacgaagvgcgaca	actvtgvgaaag	1068	
OY	1266	ccttattgtt	vgcgaaggaat	gtvgvgcgt	tgcttcgcgaagcat	tcaacactgttcagaaac	1325	
Dp	1069	ccttattgtt	vgcgaaggaat	gtvgvgcgt	tgcttcgcgaagcat	tcaacactgttcagaaac	1128	
OY	1326	aagvggacac	attcaagaaag	gaagccttaca	tttgaaaggaatgtvgagcaag	gctttagc	1385	
Dp	1129	aagvggacac	attcaagaaag	gaagccttaca	tttgaaaggaatgtvgagcaag	gctttagc	1188	
OY	1386	cagaagcttag	cgcaccccttaca	aaaccccttaag	gacacacacggvgaaagccttattgtagc		1445	
Dp	1189	cagaagcttag	cgcaccccttaca	aaaccccttaag	gacacacacggvgaaagccttattgtagc		1248	
OY	1446	acgaagatg	vgvgcgtcaca	ctttagctg	gaaatcaaaaccc	tacaacacccgaagaaac	1505	
Dp	1249	acgaagatg	vgvgcgtcaca	ctttagctg	gaaatcaaaaccc	tacaacacccgaagaaac	1308	
OY	1506	tcaagvggttaa	accttattgt	ctcgtcgt	gaggtgcggaagtgctttagcctgaag	tcaaac	1555	
Dp	1309	tcaagvggttaa	accttattgt	ctcgtcgt	gaggtgcggaagtgctttagcctgaag	tcaaac	1368	
OY	1566	cttacaacaa	acacccagvg	gcacacac	ccggvggaagc	catcttgtattatcgaagatg	vg	1624
Dp	1369	cttacaacaa	acacccagvg	gcacacac	ccggvggaagc	catcttgtattatcgaagatg	vg	1428
OY	1625	gcgaagctt	taacccggaaatca	aacccctg	taacgcgaagaaagaaac	actcagvggaagaa	1684	
Dp	1429	gcgaagctt	taacccggaaatca	aacccctg	taacgcgaagaaagaaac	actcagvggaagaa	1488	
OY	1685	gcaatctgtat	gtgtgtgag	tgtgacgaag	gctttagatgaag	tcaaccccttaca	1744	
Dp	1489	gcaatctgtat	gtgtgtgag	tgtgacgaag	gctttagatgaag	tcaaccccttaca	1548	
OY	1745	ccgaagagaca	catcacaagvgg	aaaaagcctt	ttagtvgcgaagvg	tgtvgcgaagaaag	ctttgc	1804
Dp	1549	ccgaagagaca	catcacaagvgg	aaaaagcctt	ttagtvgcgaagvg	tgtvgcgaagaaag	ctttgc	1608
OY	1805	gcaagaaac	cttaaacctgt	tttaggcaca	agvgggaacac	ctagaatg	tgcttvtg	1844
Dp	1609	gcaagaaac	cttaaacctgt	tttaggcaca	agvgggaacac	ctagaatg	tgcttvtg	1688
OY	1865	gtgtgtg	gcaag	gctttgtgt	cttagttaa	ctcacaacacgaagacac	gcaag	1924
Dp	1669	gtgtgtg	gcaag	gctttgtgt	cttagttaa	ctcacaacacgaagacac	gcaag	1728
OY	1925	ggggaaac	cttacaat	tgtvgcgaag	gaaggtvgvgcgaag	ctttagcggcaag	ctacat	1984
Dp	1729	ggggaaac	cttacaat	tgtvgcgaag	gaaggtvgvgcgaag	ctttagcggcaag	ctacat	1788
OY	1985	tagacac	ccagagacac	atcaca	ttcagaagaaag	tgtvgcagvg	2044	

Df	1799	tagaacaccagaggacaccttcagggagaagccctatatttgcagaaagtgtggaaccggg	1848
Oy	2045	cttttgtcgggaagtcaccaacttatcacagcatcaagacaactcagataaaaactta	2104
Dd	1849	ccttagtgcggaagtcgaaccttacatcagacalcagagagacaeactcagatatgaactta	1908
Oy	2105	tgtgatatagggaalgtgtgtlacagaccttagccaagagatcatacttcatacagaccagag	2164
Dd	1909	tgtgtatagggaatgtgtgtacagacctttagcccagagatcatacttcatacagaccagag	1968
Oy	2165	gacacacacagtgctgtgtgcttttccagccattgctaataccaagtgsagacattcy	2224
Dd	1969	gacacacacacagtgctgtgtgcttttccagccattgctaataccaagtgsagacattcy	2028
Oy	2225	tgtgtatattatgcatacaga	2243
Dd	2029	tgtgtatattatgcatacaga	2047
<hr/>			
RESULT 2			
ID	AAF44909	standard; cDNA: 678 BP.	
XX	AAAF4909;		
AC			
XX			
DT	28-MAR-2001	(first entry)	
XX			
DE		Human breast cancer related protein coding sequence SEQ ID NO: 65.	
KX		Human; breast cancer; diagnosis; therapy; vaccine; ss.	
RW		Homo sapiens.	
OS			
XX			
PN	WO20078960-A2.		
XX			
PD	28-DEC-2000.		
XX			
PF	23-JUN-2000; 2000WO-US17536.		
XX			
PR	23-JUN-1999; 99US-0140903.		
PR	12-OCT-1999; 99US-0158980.		
XX			
PA	(CORI-) CORIXA CORP.		
XX			
PI	Yugu J, Mitcham JL;		
XX			
DR	WPJ; 2001-041426/05.		
XX			
PT	New polynucleotides encoding breast tumor specific proteins, useful for prevention, treatment and diagnosis of breast cancer -		
XX			
PS	Claim 25; Page 141; 165pp; English.		
CC	The present invention provides the coding sequences for a number of breast cancer related proteins. These can be used in vaccinations against, diagnosis of and treatment of cancer, particularly breast cancer.		
CC			
XX			
SQ	Sequence 678 BP; 165 A; 167 C; 160 G; 182 T; 4 other:		
Query Match 20.6%; Score 571; DB 22; Length 678; Best Local Similarity 100.0%; Pred. No. 7.2e-271; Indels 0; Gaps 0. Matches 571; Conservative 0; Mismatches 0;			
Oy	2078	gaggacacactcaggaatagaacttatbtgtatagggaatglgttacagccttagcca	2137
Dd	9	gaggacacactcaggaatagaactttaatgtgtatagggaatglgttacagccttagcca	68
Oy	2138	ggagtcatacttcatacagacaccagagagacacacagatgctgtgtcttttcaaaccatt	2197
Dd	69	ggagtcatacttcatacagacaccagagagacacacacagatgctgtgtcttttcaaaccatt	128

QY	2198	gctatgataccaagaatgtagacacattcgtgtgtgattatgcatgtagagactgtactggttaaga	2257
Db	129	gctatgataccaagaatgtagacacattcgtgtgtgattatgcatgtagagactgtactggttaaga	188
QY	2258	ctgtatctcatccatccacttaagaagaatgtcgtgcctcatcttccaggaagccctgccttc	2317
Db	189	ctgtatctcatccatccacttaagaagaatgtcgtgcctcatcttccaggaagccctgccttc	248
QY	2318	ctcacgtgtgatatgtgtgtgtgtgaaaccgcgtcagtgatagtatgtgtgcaagagagcagt	2377
Db	249	ctcacgtgtgatatgtgtgtgtgtgaaaccgcgtcagtgatagtatgtgtgcaagagagcagt	308
QY	2378	caaatgtcccgcgcgagatagaagggtggtgtacctgtgtgaaccaccaacttaagaagtgaagaca	2437
Db	309	caaatgtcccgcgcgagatagaagggtggtgtacctgtgtgaaccaccaacttaagaagtgaagaca	368
QY	2438	gtcccgcgctaatactcatatactgaatttgagaacctgtcttcccatattgtgtgtcttctc	2497
Db	369	gtcccgcgctaatactcatatactgaatttgagaacctgtcttcccatattgtgtgtcttctc	428
QY	2498	ccgattatcccaaaccttaccatatttaagatataccgcgccttctcctaattgtttt	2557
Db	429	ccgattatcccaaaccttaccatatttaagatataccgcgccttctcctaattgtttt	488
QY	2558	acaatgtgtgtgcacacctttgtgtgtgtgccttgcatacttacaatcaatgaatcaacggt	2617
Db	489	acaatgtgtgtgcacacctttgtgtgtgtgccttgcatacttacaatcaatgaatcaacggt	548
QY	2618	attccctcatcttgagccataaagaacca	2648
Db	549	attccctcatcttgagccataaagaacca	579

XX	RESULT 3
XX	AAS68256
ID	AAS68256 standard; cDNA; 653 BP.
XX	
AC	AAS68256;
DT	13-FEB-2002 (first entry)
XX	
DE	DNA encoding novel human diagnostic protein #4060.
DE	
XX	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss
XX	
OS	Homo sapiens.
XX	
PN	WO200175067-A2.
XX	
PD	11-OCT-2001.
XX	
PE	30-MAR-2001; 2001MO-US08631.
PR	31-MAR-2000; 2000US-0540217.
PR	23-AUG-2000; 2000US-0649167.
XX	
PA	(HYSE-) HYSEQ INC.
DE	
PI	Drmanac RT, Liu C, Tang YT;
XX	
DR	WPI; 2001-639362/73.
DR	P-PsDB; ABG04069.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity -
PS	Claim 1; SEQ ID No 4060; 103bp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and

polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pcl_sequences](http://wipo.int/pub/published_pcl_sequences).

Sequence 653 BP; 192 A; 151 C; 176 G; 134 T; 0 other;

Query Match	18.98;	Score 525;	DB 23;	Length 653;
Best Local Similarity	99.88;	Pred. No. 3.3e-248;		
Matches 575;	Conservative	1;	Indels 0;	Gaps 0;

OY	1004	tcgggggaaagccttatgtgtgtcagaggaatgtgtggtcgagagccttacgttgcgtggaattcaacct	1063
Db	52	tcgggggaaagccttatgtgtgtcagaggaatgtgtggtcgagagccttacgttgcgtggaattcaacct	111
OY	1064	gacccacatcagaagagacacactcgaagggaagagaacctatgtgtcagaagattgtggagc	1123
Db	112	gacccacatcagaagagacacactcgaagggaagagaacctatgtgtcagaagattgtggagc	171
OY	1124	aggcttacttggaaagtcgaacctcttaccatcagcggagacacactcaaggctcaagcc	1183
Db	172	aggcttacttggaaagtcgaacctcttaccatcagcggagacacactcaaggctcaagct	231
OY	1184	ttatgtgtgcagaaggaattgtggcagaagctttagccttgtaagtcaaacctcatcaccaca	1243
Db	232	ttatgtgtgcagaaggaattgtggcagaagctttagccttgtaagtcaaacctcatcaccaca	291
OY	1244	gagggcgacacactgggggagaagccttatgttttcacgggagatgtggcgctgtgttcgcca	1303
Db	292	gagggcgacacactgggggagaagccttatgttttcacgggagatgtggcgctgtgttcgcca	351
OY	1304	gcattcacacctgtgtcagacacaaagaagacacatctcagaagagaagaagccttacattgcag	1363
Db	352	gcattcacacctgtgtcagacacaaagaagacacatctcagaagagaagaagccttacattgcag	411
OY	1364	ggaagtgtgagcagaagcctttagccagaagaatcaacctctcagacaccttaagaacacaaac	1423
Db	412	ggaagtgtgagcagaagcctttagccagaagaatcaacctctcagacaccttaagaacacaaac	471
OY	1424	aggagagaagccttatgttatgcacagaagtgtgtgggtcactttagcttgcgaaatcaacct	1483
Db	472	aggagagaagccttatgttatgcacagaagtgtgtgggtcactttagcttgcgaaatcaacct	531
OY	1484	caaaacacacacagaagacacactcaggggtttaaacttatgtctgcctgtggaattggggca	1543
Db	532	caaaacacacacagaagacacactcaggggtttaaacttatgtctgcctgtggaattggggca	591
OY	1544	gtgccttagcctgaagtcaaaccttaacaaacaca	1579
Db	592	gtgccttagcctgaagtcaaaccttaacaaacaca	627

RESULT	4
AAS26286	
ID	AAS26286 standard; cDNA; 983 BP.
XX	

AC AAS26286;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, Seq ID 465.
XX
KW Human: immunosuppressive; antiarthritic; ss; antirheumatic;
KW cytosolic; cardiant; vasotropic; cerebroprotective; nocotropic;
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
KW vulnery; secreted protein; rheumatoid arthritis;
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
KW cerebrovascular disorder; cerebral ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
KW corneal infection; wound healing; epithelial cell proliferation;
KW skin ageing; food additive; preservative; antiproliferative.
XX
OS Homo sapiens.
XX
PN WO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01341.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232402.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 13-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241096.
PR 20-OCT-2000; 2000US-0241121.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.

17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-488783/53.
XX P-PSDB; AAU16299.

PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -

PS Claim 1; SEQ ID NO 465; 980bp; English.

XX The invention relates to isolated nucleic acid molecules and their
XX encoded secreted proteins. The nucleic acids and proteins are used to
XX prevent, treat or ameliorate a medical condition in e.g. humans, mice,
XX rabbits, goats, horses, cats, dogs, chickens or sheep. They
XX are also used in diagnosing a pathological condition or susceptibility
XX to a pathological condition. Antibodies to the proteins can also
XX be used in alleviating symptoms associated with the disorders and in
XX diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
XX immunosorbent assays (ELISA). Disorders which are diagnosed or treated
XX include autoimmune diseases e.g. Rheumatoid arthritis,
XX hyperproliferative disorders e.g. neoplasms of the breast or liver,
XX cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
XX e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
XX Alzheimer's disease, infections caused by bacteria, viruses and fungi
XX and ocular disorders e.g. corneal infection. The polypeptides can also
XX be used to aid wound healing and epithelial cell proliferation, to
XX prevent skin aging due to sunburn, to maintain organs before
XX transplantation, for supporting cell culture of primary tissues, to
XX regenerate tissues and in chemotaxis. The polypeptides can also be used
XX as a food additive or preservative to increase or decrease storage
XX capabilities, fat content, lipid, protein, carbohydrate, vitamins,
XX minerals, cofactors and other nutritional components. The present
XX sequence encodes a novel secreted protein of the invention.

Query Match 8.3%; Score 231; DB 22; Length 983;

Best Local Similarity 99.6%; Pred. No. 2.4e-103; Mismatches 0;

Matches 281; Conservative 0; Indels 0; Gaps 0;

Db 287 gaacggaggagtgatgctgagactataacatctggtctacagtggaatccatctc 346
|||
Db 570 gaacggaggagtgatgctgagactataacatctggtctacagtggaatccatctc 629
|||
Oy 347 taacccaaactcattgtctcagctgagcgagggagagccctctgagagagagaa 406
|||
Db 630 taacccaaactcattgtctcagctgagcgagggagagccctctgagagagagaa 689
|||
Oy 407 atgtccactgagccctctgctcgaatcgagccagaatcaactagtcctctgccc 466
|||
Db 690 atgtccactgagccctctgctcgaatcgagccagaatcaactagtcctctgccc 749
|||
Oy 467 tctgatttctcctcagtcagcagctctcagccaacatgtctgagtcagtcactctca 526
|||
Db 750 tctgatttctcctcagtcagcagctctcagccaacatgtctgagtcagtcactctca 809

Oy 527 gctgtttcagaattatgggcaggaatcctctccaccctggg 568
|||
Db 810 gctgtttcagaattatgggcaggaatcctctccaccctggg 851

RESULT 5

AAI83890 standard; cDNA; 394 BP.

AAI83890;

06-NOV-2001 (first entry)

Human polynucleotide SEQ ID NO 3950.

Human; cytokine; cell proliferation; cell differentiation; gene therapy;
vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
tissue growth factor; immunomodulatory; cancer; leukaemia;
nervous system disorders; arthritis; inflammation; ss.

Homo sapiens.

WO200164835-A2.

07-SEP-2001.

26-FEB-2001; 2001WO-US04927.

28-FEB-2000; 2000US-0515126.

18-MAY-2000; 2000US-0577409.

(HYSE-) HYSEQ INC.

Tang YF, Liu C, Drmanac RT;

WPI; 2001-514838/56.

P-PSDB; AA003959.

Isolated nucleic acids and polypeptides, useful for preventing
diagnosing and treating e.g. leukaemia, inflammation and immune
disorders -

Claim 1; SEQ ID NO 3950; 1399bp + Sequence Listing; English.

The invention relates to human polynucleotides (AAI79941-AAI93841) and
the encoded proteins (AA000010-AA013910) that exhibit activity elating to
cytokine, cell proliferation or cell differentiation or which may induce
production of other cytokines in other cell populations. The
polynucleotides and polypeptides are useful in gene therapy, vaccines or
peptide therapy. The polypeptides have various cytokine-like activities,
e.g. stem cell growth factor activity, haematopoiesis regulating
activity, tissue growth factor activity, immunomodulatory activity and
activin/inhibin activity and may be useful in the diagnosis and/or
treatment of cancer, leukaemia, nervous system disorders, arthritis and
inflammation.
Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from Wipo
at ftp.wipo.int/pub/published_pcl_sequences.

Sequence 394 BP; 148 A; 75 C; 93 G; 78 T; 0 other;

Query Match 3.6%; Score 99; DB 22; Length 394;

Best Local Similarity 99.3%; Pred. No. 2.6e-38; Mismatches 1; Indels 0; Gaps 0;

Matches 149; Conservative 0; Indels 0; Gaps 0;

Oy 2619 ttcctcattctgagccataaaagaccagactcagctgagtgagagagaatcacc 2678
|||
Db 1 ttcctcattctgagccataaaagaccagactcagctgagtgagagagaatcacc 60
|||
Oy 2679 ctgctgtgagggttgggagaccctcctcgtgacccctcctcagctgagtgagtggtttcttt 2738
|||
Db 61 ctgctgtgagggttgggagaccctcctcgtgacccctcctcagctgagtgagtggtttctttg 120

OY 2739 ccaataaattctttcaccatccca 2768
|||||
DB 121 cccaataaattctttcaccatccca 150

RESULT 6
AAS13031

ID AAS13031 standard: DNA; 69 BP.

AC AAS13031;

DT 17-DEC-2001 (first entry)

DE DNA encoding zinc finger domain TG-ZFD-012.

KM Zinc finger domain; cancer; human; ds: TG-ZFD-012.

OS Homo sapiens.

PN W0200160970-A2.

PD 23-AUG-2001.

PF 17-FEB-2001; 2001WO-KR00244.

PR 18-FEB-2000; 2000KR-0007730.

PA (TOOL-) TOOLGEN INC.

PI Kim J, Kwon Y, Kim H, Ryu EH, Hwang MS;

DR WPI: 2001-557644/62.

DR P-PSDB: AAU08514.

PT Identifying a zinc finger domain for e.g. designing new polypeptides
that bind to a specific site on a DNA, comprises expressing hybrid
nucleic acids with a test zinc finger domain in cells -

PS Example 22; Page 55; 147pp; English.

CC The invention relates to a method of identifying a zinc finger domain
CC that recognises a target site on a DNA. The method comprises expressing
CC hybrid nucleic acids with a test zinc finger domain in cells containing a
CC reporter construct, where the reporter gene is expressed above a given
CC level when a transcription factor recognises a recruitment and a target
CC site of a promoter, and not only the recruitment site of the promoter.
CC The method is used to: (a) identify a zinc finger domain that recognises
CC a target site on a DNA; (b) determine whether a test zinc finger domain
CC recognises a target site on a promoter; (c) generate a nucleic acid that
CC encodes a chimeric zinc finger protein; and (d) identify DNA sequences
CC recognised by zinc finger domains. The method can be used to design novel
CC polypeptides that bind to a specific site on a DNA. The method can
CC facilitate the customised generation of new polypeptides that can
CC regulate the expression of a selected target e.g. a gene required by a
CC pathogen can be repressed, a gene required for cancerous growth can be
CC repressed, or a gene poorly expressed or encoding a mutated protein can
CC be activated and overexpressed. The method can be used in vivo which
CC enables identification of polypeptides that bind to a specific site on a
CC DNA in the intracellular milieu. The present sequence represents the
CC coding sequence of TG-ZFD-012 zinc finger protein which was used in
CC the method of the invention.

SQ Sequence 69 BP; 17 A; 16 C; 20 G; 16 T; 0 other;

Query Match 2.5%; Score 69; DB 22; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1269 tatgttcaggaagtgtgagcgttcgcgaacatccacactgtgcagacaag 1328
|||||
DB 1 tatgttcaggaagtgtgagcgttcgcgaacatccacactgtgcagacaag 60

OY 1329 aggaacacat 1337
|||||
DB 61 aggaacacat 69

RESULT 7
AAS13088

ID AAS13088 standard: DNA; 69 BP.

AC AAS13088;

DT 17-DEC-2001 (first entry)

DE DNA encoding zinc finger domain TG-ZFD-046.

KM Zinc finger domain; cancer; human; ds.

OS Homo sapiens.

PN W0200160970-A2.

PD 23-AUG-2001.

PF 17-FEB-2001; 2001WO-KR00244.

PR 18-FEB-2000; 2000KR-0007730.

PA (TOOL-) TOOLGEN INC.

PI Kim J, Kwon Y, Kim H, Ryu EH, Hwang MS;

DR WPI: 2001-557644/62.

DR P-PSDB: AAU08558.

PT Identifying a zinc finger domain for e.g. designing new polypeptides
that bind to a specific site on a DNA, comprises expressing hybrid
nucleic acids with a test zinc finger domain in cells -

PS Example 54; Page 69; 147pp; English.

CC The invention relates to a method of identifying a zinc finger domain
CC that recognises a target site on a DNA. The method comprises expressing
CC hybrid nucleic acids with a test zinc finger domain in cells containing a
CC reporter construct, where the reporter gene is expressed above a given
CC level when a transcription factor recognises a recruitment and a target
CC site of a promoter, and not only the recruitment site of the promoter.
CC The method is used to: (a) identify a zinc finger domain that recognises
CC a target site on a DNA; (b) determine whether a test zinc finger domain
CC recognises a target site on a promoter; (c) generate a nucleic acid that
CC encodes a chimeric zinc finger protein; and (d) identify DNA sequences
CC recognised by zinc finger domains. The method can be used to design novel
CC polypeptides that bind to a specific site on a DNA. The method can
CC facilitate the customised generation of new polypeptides that can
CC regulate the expression of a selected target e.g. a gene required by a
CC pathogen can be repressed, a gene required for cancerous growth can be
CC repressed, or a gene poorly expressed or encoding a mutated protein can
CC be activated and overexpressed. The method can be used in vivo which
CC enables identification of polypeptides that bind to a specific site on a
CC DNA in the intracellular milieu. The present sequence represents the
CC coding sequence of TG-ZFD-046 zinc finger protein which was used in
CC the method of the invention.

SQ Sequence 69 BP; 21 A; 14 C; 18 G; 16 T; 0 other;

Query Match 2.5%; Score 69; DB 22; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.6e-23;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2019 tatatttcaggaagtgtgagcgttcagtcggaagtccacactatcacacatcaag 2078
|||||
DB 1 tatatttcaggaagtgtgagcgttcagtcggaagtccacactatcacacatcaag 60

QY 2079 aggcacac 2087
|||||||
Db 61 aggcacac 69

RESULT 8
AA568257
ID AA568257 standard; cDNA; 821 BP.
XX
AC AA568257;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #4061.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR P-PSDB; ABG04070.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 1; SEQ ID No 4061; 103bp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridization probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging or sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations in
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AA564197-AA594364 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 821 BP; 174 A; 234 C; 215 G; 198 T; 0 other;

Query Match 1.8%; Score 49; DB 23; Length 821;
Best Local Similarity 100.0%; Pred. No. 1.1e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 722 acacagccagcacagtcacagggaacacacacagctggtgatataagg 770
|||||||
Db 188 acacagccagcacagtcacagggaacacacacagctggtgatataagg 236

RESULT 9
AAH33832
ID AAH33832 standard; cDNA; 800 BP.
XX
AC AAH33832;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:888.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; ss.
XX
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI: 2001-235357/24.
DR P-PSDB; AAG74401.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
PT
XX
PS Claim 1; Page 2813-2814; 9803bp; English.
XX
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated cell
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAG77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 800 BP; 240 A; 173 C; 219 G; 166 T; 2 other;

Query Match 1.0%; Score 29; DB 22; Length 800;
Best Local Similarity 100.0%; Pred. No. 0.00079;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1659 caccagagacacactcaggaggaagcc 1687
|||||||
Db 413 caccagagacacactcaggaggaagcc 441

RESULT 10
 AAS90429
 ID AAS90429 standard; CDNA: 2349 BP.
 XX
 AC AAS90429;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #26233.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HXSE-) HXSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR P-PSDB; ABG26242.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 XX
 PS Claim 1; SEQ ID NO 26233; 103pp; English.
 XX
 XX
 XX "The invention relates to isolated polynucleotide (I) and
 XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
 XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 XX and gene mapping, and in recombinant production of (II). The
 XX polynucleotides are also used in diagnostics as expressed sequence tags
 XX for identifying expressed genes. (I) is useful in gene therapy techniques
 XX to restore normal activity of (II) or to treat disease states involving
 XX (II). (II) is useful for generating antibodies against it, detecting or
 XX quantitating a polypeptide in tissue, as molecular weight markers and as
 XX a food supplement. (II) and its binding partners are useful in medical
 XX imaging of sites expressing (II). (I) and (II) are useful for treating
 XX disorders involving aberrant protein expression or biological activity.
 XX The polypeptide and polynucleotide sequences have applications in
 XX diagnostics, forensics, gene mapping, identification of mutations
 XX responsible for genetic disorders or other traits to assess biodiversity
 XX and to produce other types of data and products dependent on DNA and
 XX amino acid sequences. AAS64197-AAS94564 represent novel human
 XX diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 CC
 CC Sequence 2349 BP; 664 A; 547 C; 608 G; 530 T; 0 other;
 CC

Query Match	1.0%	Score 29	DB 23	Length 2349
Best Local Similarity	100.0%	Pred. No.	0.00079	
Matches 29	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Qy	1659	caccagaagacactcaggaggagaagcc	1687
Db	1645	caccagaagacacactcaggaggagaagcc	1673

RESULT 11

AA158664
ID AA158664 standard; cDNA; 3582 BP.
AC AA158664;
XX
XX
DT 22-OCT-2001 (first entry)
XX
XX DE Human polynucleotide SEQ ID NO 867.
XX
XX Human; noctropic; immunosuppressant; cytosolic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukemia; ss.
XX
XX Homo sapiens.
OS
XX
XX WO20015312-A1.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 26-DEC-2000; 2000WO-US34263.
PF
XX
XX 21-JAN-2000; 2000US-0488725.
PR
XX 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-063450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HSE-) HXSEQ INC.
PA
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
PI
XX WPI: 2001-442253/47.
DR P-PSDB; AAM39508.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
PT
XX
XX Claim 1; SEQ ID NO 867; 10078pp; English.
PS
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with noctropic,
CC immunosuppressant and cytosolic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokine activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
XX specification.
XX
XX Sequence 3582 BP; 1024 A; 826 C; 910 G; 821 T; 1 other:

Query Match	1.0;	Score 29;	DB 22;	Length 3582;
Best Local Similarity	100.0%;	Pred. No. 0.00078;		
Matches 29; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1659 caccagagcacactcagggaagcc 1687
|||||

DB 1487 caccagagacacactcagggaagcc 1515

RESULT 12

AAI60450/C

ID AAI60450 standard; cDNA; 3582 BP.

XX

AC AAI60450;

XX

DT 22-OCT-2001 (first entry)

XX

DE Human polynucleotide SEQ ID NO 4439.

XX

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KW chemokine; thrombolytic; drug screening; arthritis; inflammation;

KW Leukemia; ss.

XX

OS Homo sapiens.

XX

PN WO200153312-A1.

XX

PD 26-JUL-2001.

XX

PF 26-DEC-2000; 2000MO-US34263.

XX

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX

DR WFI; 2001-442253/47.

XX

DR P-PSDB; AAM41294.

XX

PT Novel nucleic acids and polypeptides, useful for treating disorders

PT such as central nervous system injuries -

XX

PS Claim 1; SEQ ID NO 4439; 10078bp; English.

XX

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and

CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: Immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and

CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed

CC specification.

XX

SQ Sequence 3582 BP; 821 A; 908 C; 826 G; 1026 T; 1 other;

Query Match 1.0%; Score 29; DB 22; Length 3582;

Best Local Similarity 100.0%; Pred. No. 0.00078;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1659 caccagagacacactcagggaagcc 1687

DB 2096 CACCAGAGACACTCAGGAGAAACC 2068

RESULT 13

AAL00216

ID AAL00216 standard; cDNA; 159 BP.

XX

AC AAL00216;

XX

DT 21-NOV-2001 (first entry)

XX

DE Human reproductive system related antigen cDNA SEQ ID NO: 217.

XX

KW Human; reproductive system related antigen; reproductive system disorder;

KW cancer; gene therapy; ss.

XX

OS Homo sapiens.

XX

PN WO200155320-A2.

XX

PD 02-AUG-2001.

XX

PF 17-JAN-2001; 2001WO-US01339.

XX

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226686.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

```
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235835.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.

PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249269.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250161.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251031.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI: 2001-465570/50.
DR P-PSDB; AAM94246.
XX
XX Isolated nucleic acid molecule encoding a reproductive system antigen
PT is used in preventing, treating or ameliorating a medical condition -
XX
XX Claim 1; SEQ ID NO 217; 1297pp + Sequence listing; English.
PS
XX
XX The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a coding sequence of the
CC invention.
XX
XX
SQ Sequence 159 BP; 47 A; 33 C; 47 G; 30 T; 2 other;

Query Match 1.0%; Score 28; DB 22; Length 159;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1659 caccagagagacactcagggagaagc 1686
Db 18 caccagagagacacaccagggagaagc 45

RESULT 14
AAS25949
ID AAS25949 standard; cDNA; 1183 BP.
XX
AC AAS25949;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, Seq ID 128.
XX
XX Human; immunosuppressive; antiarthritic; ss; antirheumatic;
KW cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
KW vulnery; secreted protein; rheumatoid arthritis;
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
KW cerebrovascular disorder; cerebral ischemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
KW corneal infection; wound healing; epithelial cell proliferation;
KW skin ageing; food additive; preservative; antiproliferative.
XX
OS Homo sapiens.
XX
XX WO200155322-A2.
PN
XX
```

PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01341.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0255678.

(HUMA-) HUMAN GENOME SCI INC.
PA
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI, 2001-488783/53.

PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-488783/53.
XX
XX New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -
XX
XX
PS Disclosure; SEQ ID No 1609; 980pp; English.
XX

CC The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
CC are also used in diagnosing a pathological condition or susceptibility
CC to a pathological condition. Antibodies to the proteins can also
CC be used in alleviating symptoms associated with the disorders and in
CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
CC immunosorbant assays (ELISA). Disorders which are diagnosed or treated
CC include autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
CC and ocular disorders e.g. corneal infection, and many other
CC disorders listed in the specification. The polypeptides can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence is a genomic DNA encoding a partial novel secreted protein of
CC the invention.

Query Match 1.0%; Score 27; DB 22; Length 2717;
Best Local Similarity 100.0%; Pred. No. 0.0076;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2149 tcattcagacacagagacacacag 2175
Db 962 tcattcagacacacagagacacacag 988

RESULT 16
AAV64361
ID AAV64361 standard; cDNA; 3099 BP.
XX
XX AAV64361;
AC
XX
DT 15-FEB-1999 (first entry)
XX
XX Human stem cell zinc finger SZFL-2 cDNA.
DE
XX
XX SZFL-2; human; stem cell zinc finger; transcription factor;
KM haematopoiesis; CD34+; gene therapy; ss.
XX
XX Homo sapiens.
OS
XX
XX
FH Key Location/Qualifiers
FT CDS 379..1465
FT /*tag= a
XX
XX WO9845326-A1.
XX
XX 15-OCT-1998.
XX
XX PD
XX PF 08-APR-1998; 98WO-US06925.
XX
XX PR 08-APR-1997; 97US-0041811.
XX
XX PA (OSIR-) OSIRIS THERAPEUTICS INC.
XX
XX
XX Civin CI, Liu C, Small D;
XX
XX WPI; 1999-009328/01.
DR P-PSDB; AAWB1632.
XX
XX New polynucleotides encoding stem cell zinc fingers - useful for
PT gene therapy to replace or supplement defective gene for those
PT proteins
XX

PS Disclosure: Page 44-46; 76pp; English.

XX This cDNA sequence includes a coding region for human stem cell
CC zinc finger protein SZFI-2 (see AAW81632), a novel transcription
CC factor that appears to be expressed in most cell types and tissues
CC and which may be involved in the maturation of CD34+ cells; certain
CC concentrations of SZFI-2 can cause haematopoietic cells to
CC differentiate and mature rather than to merely replicate. The
CC identification of SZFI was made through the random sequencing of a
CC portion of several hundred clones from a cDNA library prepared from
CC human bone marrow CD34+ cells. One of the fragments was unique but
CC showed homology to the Kruppel family of zinc finger proteins.
CC This fragment was used to screen cDNA libraries from CD34+ cells,
CC K562 cells and human lung. Overlapping cDNA fragments which
CC hybridised to the probe gave rise to 2 alternatively spliced cDNA
CC products, termed SZFI-1 (see AAV64360) and SZFI-2. The gene is
CC found on chromosome 3. Database searches showed the most highly
CC related genes to be ZNF133, KID1 and ZNF85 (65%, 53% and 45%
CC homology at the nucleotide level, respectively). The invention
CC additionally provides recombinant vectors and host cells, and a
CC method of producing SZFI polypeptides. The polynucleotides may be
CC utilised for gene therapy in a host to replace or supplement a
CC defective SZFI gene.

XX

SQ Sequence 3099 BP; 836 A; 744 C; 812 G; 707 T; 0 other;

Query Match 1.0%; Score 27; DB 20; Length 3099;
Best Local Similarity 100.0%; Pred. No. 0.0076;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2149 tcacagaccagagacacacag 2175
DB 1913 tcacagaccagagacacacag 1939
|||||

RESULT 17
AB100474/C
ID ABL00474 standard; DNA; 51 BP.

XX ABL00474;
AC
XX
XX
DT 05-MAR-2002 (first entry)
XX
XX
DE Human silent noncoding SNP oligonucleotide SEQ ID NO:465.
XX
XX
KW Human; single nucleotide polymorphism; SNP; polymorphism; cytostatic;
KW immunosuppressive; antiinflammatory; neutroprotective; antimicrobial;
KW autoimmune disease; inflammation; cancer; nervous system disease;
KW infection; polymorphic protein; ds.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200138586-A2.
XX
XX
PD 31-MAY-2001.
XX
XX
PF 22-NOV-2000; 2000WO-US32311.
XX
XX
PR 24-NOV-1999; 990US-0167383.
XX
XX
PA (CURA-) CURAGEN CORP.
XX
XX
PI Shinkets RA, Leach M;
XX
XX
DR WPI; 2001-355949/37.
XX
XX
PT Isolated human nucleic acids comprising one or more single nucleotide
PT polymorphisms, useful for treating a subject suffering from a
PT pathology, e.g. autoimmune diseases, ascribed to the presence of a
PT sequence polymorphism -
XX
XX
PS Claim 1; Page 388; 674pp; English.

XX ABL00010 to ABL01104 represent human nucleic acid oligonucleotides
CC comprising one or more single nucleotide polymorphisms (SNPs). ABB56531
CC to ABB56903 represent human peptides encoded by some of the SNP
CC oligonucleotides. The sequences from the present invention can have
CC immunosuppressive, cytostatic, antiinflammatory, neutroprotective and
CC antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides
CC and antibodies from the present invention can be used for treating a
CC subject suffering from, at risk for, or suspected of, suffering from a
CC pathology ascribed to the presence of a sequence polymorphism. The
CC pathology may be autoimmune diseases, inflammation, cancer, diseases of
CC the nervous system, and infection by pathogenic microorganisms. The SNPs
CC are also useful for determining which forms of a characterised
CC polymorphism are present in individuals. The antibodies may be used in
CC the detection, quantitation and/or cellular or tissue localisation of a
CC polymorphic protein (e.g., for use in measuring levels of the
CC polymorphic protein within appropriate physiological samples).

XX

SQ Sequence 51 BP; 9 A; 10 C; 14 G; 18 T; 0 other;

Query Match 0.9%; Score 25; DB 23; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1378 gcttagccagaagtcacacctat 1402
DB 51 GCTTAGCCAGAGTCACACCTCAT 27
|||||

RESULT 18
AB100878/C
ID ABL00878 standard; DNA; 51 BP.

XX ABL00878;
AC
XX
XX
DT 05-MAR-2002 (first entry)
XX
XX
DE Human amino acid change SNP oligonucleotide SEQ ID NO:869.
XX
XX
KW Human; single nucleotide polymorphism; SNP; polymorphism; cytostatic;
KW immunosuppressive; antiinflammatory; neutroprotective; antimicrobial;
KW autoimmune disease; inflammation; cancer; nervous system disease;
KW infection; polymorphic protein; ds.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200138586-A2.
XX
XX
PD 31-MAY-2001.
XX
XX
PF 22-NOV-2000; 2000WO-US32311.
XX
XX
PR 24-NOV-1999; 990US-0167383.
XX
XX
PA (CURA-) CURAGEN CORP.
XX
XX
PI Shinkets RA, Leach M;
XX
XX
DR WPI; 2001-355949/37.
XX
XX
PT Isolated human nucleic acids comprising one or more single nucleotide
PT polymorphisms, useful for treating a subject suffering from a
PT pathology, e.g. autoimmune diseases, ascribed to the presence of a
PT sequence polymorphism -
XX
XX
PS Claim 1; Page 509; 674pp; English.

AB10010 to AB101104 represent human nucleic acid oligonucleotides
CC comprising one or more single nucleotide polymorphisms (SNPs). ABB56531
CC to ABB56903 represent human peptides encoded by some of the SNP
CC oligonucleotides. The sequences from the present invention can have
CC immunosuppressive, cytostatic, antiinflammatory, neutroprotective and
CC antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides
CC and antibodies from the present invention can be used for treating a
CC subject suffering from, at risk for, or suspected of, suffering from a
CC pathology ascribed to the presence of a sequence polymorphism. The
CC pathology may be autoimmune diseases, inflammation, cancer, diseases of
CC the nervous system, and infection by pathogenic microorganisms. The SNPs
CC are also useful for determining which forms of a characterised
CC polymorphism are present in individuals. The antibodies may be used in
CC the detection, quantitation and/or cellular or tissue localisation of a
CC polymorphic protein (e.g., for use in measuring levels of the
CC polymorphic protein within appropriate physiological samples).

CC antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides
CC and antibodies from the present invention can be used for treating a
CC subject suffering from, at risk for, or suspected of, suffering from a
CC pathology ascribed to the presence of a sequence polymorphism. The
CC pathology may be autoimmune diseases, inflammation, cancer, diseases of
CC the nervous system, and infection by pathogenic microorganisms. The SNPs
CC are also useful for determining which forms of a characterized
CC polymorphism are present in individuals. The antibodies may be used in
CC the detection, quantitation and/or cellular or tissue localization of a
CC polymorphic protein (e.g., for use in measuring levels of the
CC polymorphic protein within appropriate physiological samples).

XX
SQ Sequence 51 BP; 10 A; 14 C; 12 G; 15 T; 0 other;

Query Match 0.9%; Score 25; DB 23; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1362 agggagtgtgacgaagcttagcc 1386
|||||
Db 51 AGCGAGTGTGACGACGCTTAGCC 27

RESULT 19
AAK92290
ID AAK92290 standard; cDNA; 763 BP.
XX
AC AAK92290;
XX
DT .06-NOV-2001 (first entry)
XX
DE Human cDNA 5'-end sequence, SEQ ID NO: 750.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
PI EP130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI: 2001-524255/58.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
PS Claim 2; SEQ ID NO 750; 1380pp + sequence listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is the nucleotide
CC sequence of the 5'-end of a cDNA provided in the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX

SQ Sequence 763 BP; 169 A; 206 C; 217 G; 163 T; 8 other;

Query Match 0.9%; Score 25; DB 22; Length 763;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 ggatgtgctgtgtactaccag 248
|||||
Db 343 ggatgtgctgtgtactaccag 367

RESULT 20
AAK93811
ID AAK93811 standard; cDNA; 763 BP.
XX
AC AAK93811;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human cDNA clone representative sequence, SEQ ID NO: 2271.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
PI EP130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI: 2001-524255/58.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
PS Example 11; SEQ ID NO 2271; 1380pp + sequence listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence was used as the
CC representative sequence from a human clone which was used in
CC homology searches to identify the clone.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 763 BP; 169 A; 206 C; 217 G; 163 T; 8 other;

Query Match 0.9%; Score 25; DB 22; Length 763;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 ggatgtgctgtgtactaccag 248
|||||
Db 343 ggatgtgctgtgtactaccag 367

RESULT 21
AAI95339
ID AAI95339 standard; cDNA; 814 BP.
XX
AC AAI95339;
XX
DT 13-NOV-2001 (first entry)
XX
DE Human neuroblastoma expressed polynucleotide SEQ ID NO 1414.
XX
KW Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.
XX
OS Homo sapiens.
XX
PN WO20016719-A1.
XX
PD 13-SEP-2001.
XX
PF 02-MAR-2001; 2001WO-JP01629.
XX
PR 07-MAR-2000; 2000JP-0159195.
XX
PA (CHIB-) CHIBA PREFECTURE.
XX
PA (HISM) HISAMITSU PHARM CO LTD.
XX
PI Nakagawara A;
XX
DR WPI; 2001-565584/63.
XX
PT Nucleic acids originating in gene expressed in human neuroblastoma,
PT useful as probe or primer in diagnosing prognosis of human
PT neuroblastoma, malignancy and susceptibility indicator or tumour marker
PT for anti-cancer agents -
XX
PS Claim 1; Page 1061; 2979pp; Japanese.
XX
CC The invention relates to novel genes (AAI93926-AAI97963) expressed in
CC human neuroblastoma. The nucleic acids are applicable as a probe or
CC primer in diagnosing the prognosis of human neuroblastoma, malignancy and
CC susceptibility indicators or tumour markers for anti-cancer agents. The
CC gene information for diagnosing prognosis is related to factors similar
CC to that for N-myc and TrkA genes.
XX
SQ Sequence 814 BP; 240 A; 170 C; 217 G; 170 T; 17 other;

Query Match 0.9%; Score 25; DB 22; Length 814;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1416 acacacacagagagagagccttatg 1440
DB 368 acacacacagagagagagccttatg 392
|||||

RESULT 22
ABA08823
ID ABA08823 standard; cDNA; 871 BP.
XX
AC ABA08823;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human PRO1847 homologue-encoding cDNA, SEQ ID NO:599.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;

KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritis; haemostatic; antiatherosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antiulcer; ss.
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03800.
XX
PR 03-FEB-2000; 2000US-0496914.
XX
PR 27-APR-2000; 2000US-0560875.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YF, Liu C, Drmanac RT;
XX
DR WPI; 2001-457740/49.
XX
DR P-PSDB; ABB11579.
XX
PT Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX
PS Claim 1; Page 600; 1963pp; English.
XX
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides,
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (e.g., of nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention.
XX
SQ Sequence 871 BP; 162 A; 258 C; 236 G; 215 T; 0 other;

Query Match 0.9%; Score 25; DB 22; Length 871;

Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 224 ggatgtgctgtgtactaccag 248
|||||
Db 336 ggatgtgctgtgtactaccag 360

RESULT 23

AAK94575
ID AAK94575 standard; cDNA; 2200 BP.

XX AAK94575;

DT 06-NOV-2001 (first entry)

DE Human full-length cDNA, SEQ ID NO: 3494.

KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

OS Homo sapiens.

PN EPI130094-A2.

PD 05-SEP-2001.

PF 07-JUL-2000; 2000EP-0114089.

PR 08-JUL-1999; 99JP-0194486.

PR 11-JAN-2000; 2000JP-0118774.

PR 02-MAY-2000; 2000JP-0183765.

XX (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

DR WPI: 2001-524255/58.

DR P-PSDB; AAM93641.

PT 830 Primers useful for synthesizing full length cDNA clones and their

PT use in genetic manipulation -

PS Claim 8; SEQ ID NO 3494; 1380bp + sequence listing; English.

XX The invention relates to primers for synthesizing full length cDNA

CC clones. 830 cDNA molecules encoding a human protein have been

CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA

CC molecules have been determined. Primers for synthesizing the full length

CC cDNA are useful for clarifying the function of the protein encoded by

CC the cDNA. The full length clones were obtained by construction of full

CC length enriched cDNA libraries that were synthesised by the oligo-capping

CC method. The primers enable the production of the full length cDNA easily

CC without any special methods. The present sequence is a full length

CC human cDNA of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 2200 BP; 579 A; 558 C; 529 G; 534 T; 0 other;

Query Match 0.9%; Score 25; DB 22; Length 2200;

Best Local Similarity 100.0%; Pred. No. 0.074;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 224 ggatgtgctgtgtactaccag 248
|||||

Db 343 ggatgtgctgtgtactaccag 367

RESULT 24

AAK91560
ID AAK91560 standard; cDNA; 688 BP.

XX AAK91560;

DT 06-NOV-2001 (first entry)

DE Human cDNA 5'-end sequence, SEQ ID NO: 20.

KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

OS Homo sapiens.

PN EPI130094-A2.

PD 05-SEP-2001.

PF 07-JUL-2000; 2000EP-0114089.

PR 08-JUL-1999; 99JP-0194486.

PR 11-JAN-2000; 2000JP-0118774.

PR 02-MAY-2000; 2000JP-0183765.

XX (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

DR WPI: 2001-524255/58.

PT 830 Primers useful for synthesizing full length cDNA clones and their

PT use in genetic manipulation -

PS Claim 2; SEQ ID NO 20; 1380bp + sequence listing; English.

XX The invention relates to primers for synthesizing full length cDNA

CC clones. 830 cDNA molecules encoding a human protein have been

CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA

CC molecules have been determined. Primers for synthesizing the full length

CC cDNA are useful for clarifying the function of the protein encoded by

CC the cDNA. The full length clones were obtained by construction of full

CC length enriched cDNA libraries that were synthesised by the oligo-capping

CC method. The primers enable the production of the full length cDNA easily

CC without any special methods. The present sequence is the nucleotide

CC sequence of the 5'-end of a cDNA provided in the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 688 BP; 156 A; 146 C; 173 G; 210 T; 3 other;

Query Match 0.8%; Score 23; DB 22; Length 688;

Best Local Similarity 100.0%; Pred. No. 0.72;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 214 tggcattcaggatgtgctgtg 236
|||||

Db 361 tggcattcaggatgtgctgtg 383

RESULT 25

AAK93129
ID AAK93129 standard; cDNA; 688 BP.

XX AAK93129;

DT 06-NOV-2001 (first entry)

DE Human cDNA clone representative sequence, SEQ ID NO: 1589.

KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

OS Homo sapiens.

PN EPI130094-A2.

```
XX 05-SEP-2001.
PD
XX
XX 07-JUL-2000; 2000EP-0114089.
PF
XX
XX 08-JUL-1999; 99JP-0194486.
PR
XX 11-JAN-2000; 2000JP-0118774.
PR
XX 02-MAY-2000; 2000JP-0183765.
XX
XX (HELI-) HELIX RES INST.
PA
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2001-524255/58.
DR
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation.
PS
XX Example 11; SEQ ID NO 1589; 1380bp + sequence listing; English.
PS
XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence was used in
CC representative sequence from a human clone which was used in
CC homology searches to identify the clone.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
XX Sequence 688 BP; 156 A; 146 C; 173 G; 210 T; 3 other;
SQ
Query Match 0.8%; Score 23; DB 22; Length 688;
Best Local Similarity 100.0%; Pred. No. 0.72; Mismatches 0; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 214 tggcattcaggagatgtgctgtc 236
DB 361 tggcattcaggagatgtgctgtc 383
RESULT 26
ABA50438/C
ID ABA50438 standard; DNA; 1512 BP.
XX
XX ABA50438;
AC
XX
XX 01-FEB-2002 (first entry)
DT
XX
XX Human breast cell single exon nucleic acid probe #9133.
DE
XX Human; microarray; single exon probe; gene expression; breast;
KW disease; cancer; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200157271-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00662.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR
XX 26-MAY-2000; 2000US-0207456.
PR
XX 30-JUN-2000; 2000US-0608408.
PR
XX 03-AUG-2000; 2000US-0632366.
PR
XX 21-SEP-2000; 2000US-0234687.
```

```
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-496933/54.
XX
XX New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
XX breast, comprises number of single exon nucleic acid probes.
XX
XX Claim 4; SEQ ID NO 9133; 327bp + sequence listing; English.
PS
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pcr_sequences.
XX
XX Sequence 1512 BP; 258 A; 462 C; 386 G; 406 T; 0 other;
SQ
Query Match 0.8%; Score 23; DB 22; Length 1512;
Best Local Similarity 100.0%; Pred. No. 0.71; Mismatches 0; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1490 accaccagagacacactcaggag 1512
DB 983 ACACCAGAGACACACTCAGCGG 961
RESULT 27
ABA68388/C
ID ABA68388 standard; DNA; 1512 BP.
XX
XX ABA68388;
AC
XX
XX 01-FEB-2002 (first entry)
DT
XX
XX Human foetal liver single exon nucleic acid probe #16693.
DE
XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
KW
XX
XX Homo sapiens.
OS
XX
XX WO200157277-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00669.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR
XX 26-MAY-2000; 2000US-0207456.
PR
XX 30-JUN-2000; 2000US-0608408.
PR
XX 03-AUG-2000; 2000US-0632366.
PR
XX 21-SEP-2000; 2000US-0234687.
PR
XX 27-SEP-2000; 2000US-0236359.
PR
XX 04-OCT-2000; 2000GB-0024263.
```


XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human fetal liver -
PS Claim 4; SEQ ID NO 16693; 639pp + sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probes may be used for predicting,
XX measuring and displaying gene expression in samples derived from human
XX fetal liver. The present sequence is a single exon nucleic acid
XX probe of the invention.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 1512 BP; 258 A; 462 C; 386 G; 406 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1512;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1490 acaccagagacactcagg 1512
DB 983 ACACGAGGACACTCAGGG 961

RESULT 28
ID ABA35382/C
XX ABA35382 standard; DNA; 1512 BP.
XX
XX ABA35382;
XX
XX 23-JAN-2002 (first entry)
XX
XX Probe #13848 for gene expression analysis in human heart cell sample.
XX
XX Human; gene expression; heart; microarray; vascular system; probe;
XX cardiovascular disease; hypertension; cardiac arrhythmia;
XX congenital heart disease; ss.
XX
XX Homo sapiens.
XX
XX WO200157274-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00666.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX

PS Claim 4; SEQ ID NO 13848; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart. The
XX present sequence is one such probe. The probes may be used for
XX predicting, measuring and displaying gene expression in samples derived
XX from the human heart via microarrays. By measuring gene expression, the
XX probes are useful for predicting, diagnosing, grading, staging,
XX monitoring and prognosing diseases of the human heart and vascular system
XX e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
XX congenital heart disease.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 1512 BP; 258 A; 462 C; 386 G; 406 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1512;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1490 acaccagagacactcagg 1512
DB 983 ACACGAGGACACTCAGGG 961

RESULT 29
ID AAK16761/C
XX AAK16761 standard; DNA; 1512 BP.
XX
XX AAK16761;
XX
XX 05-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe SEQ ID NO: 16752.
XX
XX Human; brain expressed exon; gene expression analysis; probe;
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer; ss.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00667.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
XX Example 4; SEQ ID NO: 16752; 650pp + sequence listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX

CC epilepsy and cancers. The present sequence is one of the probes of the invention.
XX
SQ Sequence 1512 BP; 258 A; 462 C; 386 G; 406 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1512;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1490 acaccagagacactcaggag 1512
|||||
DB 983 ACACGAGGACACTCAGGGG 961

RESULT 30
AAK42537/c
ID AAK42537 standard; DNA; 1512 BP.

AC AAK42537;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed single exon probe SEQ ID NO: 17094.

KW Human; bone marrow expressed exon; gene expression analysis; probe; microarray; cancer; leukaemia; lymphoma; myeloma; ss.

OS Homo sapiens.

PN WO200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00668.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488900/53.

PT Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human bone marrow -

PS Example 4; SEQ ID NO: 17094; 658pp + Sequence listing; English.

CC The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukaemia and myeloma. The present sequence is one of

CC the probes of the invention.
SQ Sequence 1512 BP; 258 A; 462 C; 386 G; 406 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1512;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1490 acaccagagacactcaggag 1512
|||||
DB 983 ACACGAGGACACTCAGGGG 961

RESULT 31
AAI23284/c
ID AAI23284 standard; DNA; 1512 BP.

AC AAI23284;

DT 12-OCT-2001 (first entry)

DE Probe #13217 for gene expression analysis in human cervical cell sample.

KW Probe; human; microarray; gene expression; cervical epithelial cell;

KW cervical cancer; ss.

OS Homo sapiens.

PN WO200157278-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00670.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488901/53.

PT Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human cervical epithelial cells -

PS Claim 25; SEQ ID NO 13217; 487pp; English.

CC The present invention relates to human single exon nucleic acid probes

CC (SENP). The present sequence is one such probe. The SENPs are derived

CC from human HeLa cells. The SENPs can be used to produce a single exon

CC microarray, which can be used for measuring human gene expression in a

CC sample derived from human cervical epithelial cells. By measuring gene

CC expression, the probes are therefore useful in grading and/or staging

CC of diseases of the cervix, notably cervical cancer.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

CC
SQ Sequence 1512 BP; 258 A; 462 C; 386 G; 406 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1512;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1490 acaccagagacactcaggag 1512
|||||
DB 983 ACACGAGGACACTCAGGGG 961

RESULT 32

AAI48604/c
ID AAI48604 standard; DNA; 1512 BP.

AC AAI48604;

DT 17-OCT-2001 (first entry)

DE Probe #17290 used to measure gene expression in human placenta sample.

```

XX  Probe: microarray; human; placenta; antenatal diagnosis:
KM  genetic disorder; ss.
XX
XX  Homo sapiens.
PN  W0200157272-A2.
XX
XX  09-AUG-2001.
PD
XX
XX  30-JAN-2001: 2001MO-US00663.
PF
XX
XX  04-FEB-2000: 2000US-0180312.
PR
XX  26-MAY-2000: 2000US-0207456.
PR
XX  30-JUN-2000: 2000US-0608408.
PR
XX  03-AUG-2000: 2000US-0632366.
PR
XX  21-SEP-2000: 2000US-0234687.
PR
XX  27-SEP-2000: 2000US-0236359.
PR
XX  04-OCT-2000: 2000GB-0024263.
XX
XX  (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX  Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX  WPI: 2001-488897/53.
DR
XX
XX  Human genome-derived single exon nucleic acid probes useful for
PT  analyzing gene expression in human placenta -
XX
XX  Claim 25: SEQ ID No 17290; 654pp; English.
PS
XX
XX  The present invention relates to single exon nucleic acid probes (SENP).
CC  The present sequence is one such probe. The probes are useful for
CC  producing a microarray for predicting, measuring and displaying gene
CC  expression in samples derived from human placenta. The probes are useful
CC  for antenatal diagnosis of human genetic disorders.
XX
XX  Sequence 1512 BP; 258 A; 462 C; 386 G; 406 T; 0 other;
SQ

```

```

Query Match          0.8%; Score 23; DB 22; Length 1512;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY  1490 acaccagagacactcagg 1512
    ||||||||||||||||
DB  983 ACACCAAGAGCAGCACTCAGCG 961

```

```

RESULT 33
ID  AA108925/C
XX  AA108925 standard; DNA; 1512 BP.
XX
XX  AA108925;
AC
XX
XX  09-OCT-2001 (first entry)
DT
XX
XX  Probe #8916 used to measure gene expression in human breast sample.
DE
XX
XX  Probe: human; breast disease: breast cancer; development disorder; ss;
KM  inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
XX  Homo sapiens.
OS
XX
XX  W0200157270-A2.
PN
XX
XX  09-AUG-2001.
PD
XX
XX  29-JAN-2001: 2001MO-US00661.
PF
XX
XX  04-FEB-2000: 2000US-0180312.
PR
XX  26-MAY-2000: 2000US-0207456.
PR
XX  30-JUN-2000: 2000US-0608408.
PR

```

```

PR  03-AUG-2000: 2000US-0632366.
PR
XX  21-SEP-2000: 2000US-0234687.
PR
XX  27-SEP-2000: 2000US-0236359.
PR
XX  04-OCT-2000: 2000GB-0024263.
XX
XX  (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX  Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX  WPI: 2001-476286/51.
DR
XX
XX  Novel single exon nucleic acid probe used to measuring gene expression
PT  in a human breast -
XX
XX  Claim 25: SEQ ID No 8916; 322pp; English.
PS
XX
XX  The present invention relates to novel single exon nucleic acid probes.
CC  The present sequence is one such probe. The probes are useful for
CC  measuring human gene expression in a human breast sample, where the probe
CC  hybridises at high stringency to a nucleic acid expressed in the human
CC  breast. The probes are useful for predicting, diagnosing, grading,
CC  staging, monitoring and prognosing diseases of the human breast,
CC  particularly those diseases with polygenic aetiology. The diseases
CC  include: breast cancer, disorders of development, inflammatory diseases
CC  of the breast, fibrocystic changes, proliferative breast disease and
CC  non-carcinoma tumours.
CC  Note: The sequence data for this patent did not form part of the printed
CC  specification, but was obtained in electronic format directly from WIPO
CC  at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX  Sequence 1512 BP; 258 A; 462 C; 386 G; 406 T; 0 other;
SQ

```

```

Query Match          0.8%; Score 23; DB 22; Length 1512;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY  1490 acaccagagacactcagg 1512
    ||||||||||||||||
DB  983 ACACCAAGAGCAGCACTCAGCG 961

```

```

RESULT 34
ID  AAK94635
XX  AAK94635 standard; CDNA; 1623 BP.
XX
XX  AAK94635;
AC
XX
XX  06-NOV-2001 (first entry)
DT
XX
XX  Human full-length CDNA, SEQ ID NO: 3609.
DE
XX
XX  Human; full length CDNA; CDNA synthesis; oligo-capping; ss.
KM
XX
XX  Homo sapiens.
OS
XX
XX  EP1130094-A2.
PN
XX
XX  05-SEP-2001.
PD
XX
XX  07-JUL-2000: 2000EP-0114089.
PF
XX
XX  08-JUL-1999: 99JP-0194486.
PR
XX  11-JAN-2000: 2000JP-0118774.
PR
XX  02-MAY-2000: 2000JP-0183765.
XX
XX  (HELI-) HELIX RES INST.
PA
XX
XX  Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI  Makematsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX  WPI: 2001-524255/58.
DR
XX

```

PT 830 Primers useful for synthesizing full length cDNA clones and their
XX use in genetic manipulation -
XX
XX
PS Claim 8; SEQ ID NO 3609; 1380bp + sequence listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a full length
CC human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 1623 BP; 389 A; 346 C; 376 G; 512 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1623;
Best Local Similarity 100.0%; Pred. No. 0.71; Mismatches 0; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 214 tggcattcaggagatgtgctgtg 236
|||||
DB 361 tggcattcaggagatgtgctgtg 383

RESULT 35
ABA45305/C
ID ABA45305 standard; DNA; 1965 BP.
XX
AC ABA45305;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human breast cell single exon nucleic acid probe #4000.
XX
KW Human; microarray; single exon probe; gene expression; breast;
XX disease; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000662.
XX
PR 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
XX
PT New spatially-addressable set of single exon nucleic acid probes,
XX useful for measuring gene expression in sample derived from human
XX breast, comprises number of single exon nucleic acid probes -
XX
XX Claim 1; SEQ ID NO 4000; 327pp + sequence listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived

CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pcr_sequences.
XX
SQ Sequence 1965 BP; 366 A; 584 C; 497 G; 518 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1965;
Best Local Similarity 100.0%; Pred. No. 0.71; Mismatches 0; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1490 acaccagagagacacactcaggag 1512
|||||
DB 1110 ACACCAGAGAGACACACTCAGGGG 1088

RESULT 36
ABA55794/C
ID ABA55794 standard; DNA; 1965 BP.
XX
AC ABA55794;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #4099.
XX
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
XX
PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human foetal liver -
XX
XX Claim 1; SEQ ID NO 4099; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a single exon nucleic acid

CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1965 BP; 366 A; 584 C; 497 G; 518 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1965;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1490 acaccagagacactcaggg 1512
|||||
Db 1110 ACACGAGAGACACTCAGGG 1088

RESULT 37
ABA25474/C
ID ABA25474 standard; DNA; 1965 BP.

XX ABA25474;

DT 23-JAN-2002 (first entry)

DE Probe #3940 for gene expression analysis in human heart cell sample.

XX Human: gene expression; heart; microarray; vascular system; probe;

KW cardiovascular disease; hypertension; cardiac arrhythmia;

KW congenital heart disease; ss.

XX Homo sapiens.

PN WO200157274-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00666.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488899/53.

PT Single exon nucleic acid probes for analyzing gene expression in human

PT hearts -

PS Claim 1; SEQ ID No 3940; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for

CC measuring human gene expression in a sample derived from human heart. The

CC present sequence is one such probe. The probes may be used for

CC predicting, measuring and displaying gene expression in samples derived

CC from the human heart via microarrays. By measuring gene expression, the

CC probes are useful for predicting, diagnosing, grading, staging,

CC monitoring and prognosing diseases of the human heart and vascular system

CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and

CC congenital heart disease.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1965 BP; 366 A; 584 C; 497 G; 518 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1965;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1490 acaccagagacactcaggg 1512
|||||
Db 1110 ACACGAGAGACACTCAGGG 1088

RESULT 38
AAK04015/C
ID AAK04015 standard; DNA; 1965 BP.

XX AAK04015;

DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe SEQ ID NO: 4006.

XX Human: brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

KW epilepsy; cancer; ss.

XX Homo sapiens.

PN WO200157275-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00667.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483446/52.

PT Single exon nucleic acid probes for analyzing gene expression in human

PT brains -

PS Example 4; SEQ ID NO: 4006; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is one of the probes of the

CC invention.

XX Sequence 1965 BP; 366 A; 584 C; 497 G; 518 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1965;

Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1490 acaccagagacactcaggg 1512
|||||
Db 1110 ACACGAGAGACACTCAGGG 1088

RESULT 39
AAK29501/C

ID AAK29501 standard; DNA; 1965 BP.
XX
AC AAK29501;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 4058.
XX
KW Human: bone marrow expressed exon; gene expression analysis; probe;
microarray; cancer; leukemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PE 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 4058; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
SQ Sequence 1965 BP; 366 A; 584 C; 497 G; 518 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1965;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1490 acaccagagacacactcaggg 1512
|||||
DB 1110 ACACCAAGAGACACTCAGGG 1088

RESULT 40
AA14072/C
ID AA14072 standard; DNA; 1965 BP.
XX
AC AA14072;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #4005 for gene expression analysis in human cervical cell sample.
XX
KW Probe: human; microarray; gene expression; cervical epithelial cell;
cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.

XX
PD 09-AUG-2001.
XX
PE 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 25; SEQ ID NO 4005; 487bp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 1965 BP; 366 A; 584 C; 497 G; 518 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1965;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1490 acaccagagacacactcaggg 1512
|||||
DB 1110 ACACCAAGAGACACTCAGGG 1088

RESULT 41
AA135452/C
ID AA135452 standard; DNA; 1965 BP.
XX
AC AA135452;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #4138 used to measure gene expression in human placenta sample.
XX
KW Probe: microarray; human; placenta; antenatal diagnosis;
genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PE 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
DR
XX human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
PS Claim 25; SEQ ID No 4138; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
SQ Sequence 1965 BP; 366 A; 584 C; 497 G; 518 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1965;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1490 acaccagagacacactcagg 1512
DB 1110 ACACCAAGGACACACTCAGGG 1088

RESULT 42
AAI03925/C
ID AAI03925 standard; DNA; 1965 BP.
XX
AC AAI03925;
XX
DT 09-OCT-2001 (first entry)
XX
DE Probe #3916 used to measure gene expression in human breast sample.
XX
KW Probe: human; breast disease; breast cancer; development disorder; ss;
KW Inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US00661.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-476286/51.
DR
XX Novel single exon nucleic acid probe used to measuring gene expression
PT in a human breast -
XX
PS Claim 25; SEQ ID No 3916; 322pp; English.
XX
CC The present invention relates to novel single exon nucleic acid probes.

CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and
CC non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1965 BP; 366 A; 584 C; 497 G; 518 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 1965;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1490 acaccagagacacactcagg 1512
DB 1110 ACACCAAGGACACACTCAGGG 1088

RESULT 43
AAS81041
ID AAS81041 standard; cDNA; 2634 BP.
XX
AC AAS81041;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #16845.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR P-PsDB; ABG16854.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 16845; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting and
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAs64197-AAs94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2634 BP; 747 A; 686 C; 713 G; 488 T; 0 other;

Query Match 0.8%; Score 23; DB 23; Length 2634;
Best Local Similarity 100.0%; Pred. No. 0.71; Mismatches 0;
Matches 23; Conservative 0; Indels 0; Gaps 0;

OY 214 tggcattcaggagtgtgctgtc 236
DB 2 tggcattcaggagtgtgctgtc 24

RESULT 44

AAF67397
ID AAF67397 standard; CDNA; 439 BP.

XX AAF67397;

DT 09-APR-2001 (first entry)

XX Novel human polynucleotide, SEQ ID NO: 3159.

DE Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KM breast cancer; lung cancer; cancer detection; ss.

XX Homo sapiens.

OS WO200102568-A2.

PN 11-JAN-2001.

XX 30-JUN-2000; 2000WO-US18374.

XX 02-JUL-1999; 99US-0142310.

PR 02-JUL-1999; 99US-0142311.

XX (CHIR) CHIRON CORP.

PA (HYSE-) HYSEQ INC.

XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
PI Cirkvajakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
PI Kita D, Garcia V, Jones LM, Strache-Crain B;

XX WPI; 2001-091805/10.

PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -

XX Claim 9; Page 1017; 1046pp; English.

CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in

CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.

XX Sequence 439 BP; 104 A; 109 C; 130 G; 96 T; 0 other;

Query Match 0.8%; Score 22; DB 22; Length 439;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 228 gtgctgtgtactcaccagg 249
DB 156 gtgctgtgtactcaccagg 177

RESULT 45

AAC79234
ID AAC79234 standard; CDNA; 484 BP.

XX AAC79234;

DT 05-FEB-2001 (first entry)

XX Human lung tumour-specific CDNA #187.

XX Lung tumour protein; lung cancer; cytostatic; vaccine; ss.

OS Homo sapiens.

PN WO200060077-A2.

XX 12-OCT-2000.

XX 30-MAR-2000; 2000WO-US08560.

XX 02-APR-1999; 99US-0285323.

PR 09-AUG-1999; 99US-0370838.

PR 30-DEC-1999; 99US-0476235.

PR 03-MAR-2000; 2000US-0518809.

XX (CORI-) CORIXA CORP.

XX Reed SG, Lodes MJ, Mohamath R, Secrist H;

XX WPI; 2000-638466/61.

PT Novel lung tumor polypeptides and polynucleotides, useful for
PT detecting, monitoring or treating cancer, especially lung cancer -

XX Claim 3; Page 206; 243pp; English.

CC The present sequence is given in a specification relating to compounds
CC for therapy and diagnosis of lung cancer. Polypeptides comprising at
CC least an immunogenic part of a lung tumor protein are disclosed.
CC The polypeptides are useful for inhibiting the development of cancer,
CC especially lung cancer. Samples of T cells expressing the polypeptides
CC may be used to inhibit the development of cancer. The polypeptides are
CC also useful for detecting and monitoring the progression of cancer,
CC especially lung cancer.

XX Sequence 484 BP; 145 A; 106 C; 135 G; 98 T; 0 other;

Query Match 0.8%; Score 22; DB 21; Length 484;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1089 ggggagaaaccttatgtgtgca 1110
|||||

Db 233 ggggagaacctatgtgtgca 254

RESULT 46
ID AAD23310 standard; cDNA; 484 BP.
XX
XX AAD23310;
XX
XX 26-FEB-2002 (first entry)
XX
XX Human lung tumour-specific 2LT-120 cDNA.
DE
XX Human lung tumour protein; immunostimulant; cytostatic; gene therapy;
KW antisense-therapy; vaccine; immune response; lung cancer; 2LT-120; ss.
XX
XX Homo sapiens.
OS
XX WO200172295-A2.
PN
XX 04-OCT-2001.
PD
XX 28-MAR-2001; 2001WO-US09991.
PF
XX 29-MAR-2000; 2000US-0538037.
PR 05-JUN-2000; 2000US-0588937.
PR 18-AUG-2000; 2000US-0640878.
PR 22-SEP-2000; 2000US-234517P.
PR 01-NOV-2000; 2000US-0704512.
PR 14-DEC-2000; 2000US-0738973.
XX
XX * (CORI-) CORIXA CORP.
PA
XX Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indrias CY;
PI Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
XX
XX WPI; 2001-639201/73.
DR
XX New human lung-specific polynucleotides and polypeptides for the
PT diagnosis and treatment of disease e.g. lung cancer -
XX
XX Example 4; Page 256-257; 378pp; English.
PS
XX The invention relates to isolated lung tumour-specific proteins and
CC their corresponding cDNA molecules. Lung tumour-specific proteins and
CC their antigen-presenting cells are useful for stimulating and/or
CC expanding T cells specific for a tumour protein, and for inhibiting
CC the development of cancer. The invention also relates to a composition
CC useful for stimulating an immune response, and for treating cancer. The
CC lung tumour specific oligonucleotide is useful in gene therapy and for
CC diagnosis, detection and treatment of lung cancer. The present sequence
CC is human lung tumour-specific cDNA.
XX
XX Sequence 484 BP; 145 A; 106 C; 135 G; 98 T; 0 other;
SO

Query Match 0.8%; Score 22; DB 23; Length 484;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1089 ggggagaacctatgtgtgca 1110
|||||
Db 233 ggggagaacctatgtgtgca 254

RESULT 47
ID AAS26521 standard; cDNA; 758 BP.
XX
XX AAS26521;
AC
XX 07-NOV-2001 (first entry)
DT
XX

DE Human cDNA encoding a novel secreted protein, Seq ID 700.
XX
XX Human; immunosuppressive; antiarthritic; ss; antineumatic;
KW cyostatic; cardiant; vasotropic; cerebroprotective; nootropic;
KW neuroprotective; antibacterial; virucide; fungicide; opthalmological;
KW vulnerary; secreted protein; rheumatoid arthritis;
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
KW cerebrovascular disorder; cerebral ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
KW corneal infection; wound healing; epithelial cell proliferation;
KW skin ageing; food additive; preservative; antiproliferative.
XX
XX Homo sapiens.
OS
XX WO200155322-A2.
PN
XX 02-AUG-2001.
PD
XX 17-JAN-2001; 2001WO-US01341.
PF
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.

PR	14-SEP-2000;	2000US-0233398;
PR	14-SEP-2000;	2000US-0233399;
PR	14-SEP-2000;	2000US-0233400;
PR	14-SEP-2000;	2000US-0233401;
PR	14-SEP-2000;	2000US-0233403;
PR	14-SEP-2000;	2000US-0233464;
PR	14-SEP-2000;	2000US-0233065;
PR	21-SEP-2000;	2000US-0234223;
PR	21-SEP-2000;	2000US-0234274;
PR	21-SEP-2000;	2000US-0234997;
PR	25-SEP-2000;	2000US-0235484;
PR	25-SEP-2000;	2000US-0235488;
PR	26-SEP-2000;	2000US-0235494;
PR	27-SEP-2000;	2000US-0235834;
PR	27-SEP-2000;	2000US-0235836;
PR	29-SEP-2000;	2000US-0235327;
PR	29-SEP-2000;	2000US-0235367;
PR	29-SEP-2000;	2000US-0235368;
PR	29-SEP-2000;	2000US-0235369;
PR	29-SEP-2000;	2000US-0236370;
PR	02-OCT-2000;	2000US-0236802;
PR	02-OCT-2000;	2000US-0237037;
PR	02-OCT-2000;	2000US-0237038;
PR	02-OCT-2000;	2000US-0237039;
PR	02-OCT-2000;	2000US-0237040;
PR	13-OCT-2000;	2000US-0239355;
PR	13-OCT-2000;	2000US-0239357;
PR	20-OCT-2000;	2000US-0240960;
PR	20-OCT-2000;	2000US-0241221;
PR	20-OCT-2000;	2000US-0241785;
PR	20-OCT-2000;	2000US-0241786;
PR	20-OCT-2000;	2000US-0241787;
PR	20-OCT-2000;	2000US-0241808;
PR	20-OCT-2000;	2000US-0241809;
PR	20-OCT-2000;	2000US-0241826;
PR	01-NOV-2000;	2000US-0244617;
PR	08-NOV-2000;	2000US-0246474;
PR	08-NOV-2000;	2000US-0246475;
PR	08-NOV-2000;	2000US-0246476;
PR	08-NOV-2000;	2000US-0246477;
PR	08-NOV-2000;	2000US-0246478;
PR	08-NOV-2000;	2000US-0246523;
PR	08-NOV-2000;	2000US-0246524;
PR	08-NOV-2000;	2000US-0246525;
PR	08-NOV-2000;	2000US-0246526;
PR	08-NOV-2000;	2000US-0246527;
PR	08-NOV-2000;	2000US-0246528;
PR	08-NOV-2000;	2000US-0246532;
PR	08-NOV-2000;	2000US-0246609;
PR	08-NOV-2000;	2000US-0246611;
PR	08-NOV-2000;	2000US-0246613;
PR	17-NOV-2000;	2000US-0249207;
PR	17-NOV-2000;	2000US-0249208;
PR	17-NOV-2000;	2000US-0249209;
PR	17-NOV-2000;	2000US-0249210;
PR	17-NOV-2000;	2000US-0249211;
PR	17-NOV-2000;	2000US-0249212;
PR	17-NOV-2000;	2000US-0249213;
PR	17-NOV-2000;	2000US-0249214;
PR	17-NOV-2000;	2000US-0249215;
PR	17-NOV-2000;	2000US-0249216;
PR	17-NOV-2000;	2000US-0249217;
PR	17-NOV-2000;	2000US-0249218;
PR	17-NOV-2000;	2000US-0249219;
PR	17-NOV-2000;	2000US-0249220;
PR	17-NOV-2000;	2000US-0249221;
PR	17-NOV-2000;	2000US-0249222;
PR	17-NOV-2000;	2000US-0249223;
PR	17-NOV-2000;	2000US-0249224;
PR	17-NOV-2000;	2000US-0249225;
PR	17-NOV-2000;	2000US-0249226;
PR	17-NOV-2000;	2000US-0249227;
PR	17-NOV-2000;	2000US-0249228;
PR	17-NOV-2000;	2000US-0249229;
PR	17-NOV-2000;	2000US-0249300;
PR	01-DEC-2000;	2000US-0250160;
PR	01-DEC-2000;	2000US-0250161;
PR	05-DEC-2000;	2000US-0250391;
PR	05-DEC-2000;	2000US-0250392;

PR	05-DEC-2000;	20000US-025198B.	
PR	05-DEC-2000;	20000US-0256719.	
PR	06-DEC-2000;	20000US-0251479.	
PR	08-DEC-2000;	20000US-0251856.	
PR	08-DEC-2000;	20000US-0251868.	
PR	08-DEC-2000;	20000US-0251869.	
PR	08-DEC-2000;	20000US-0251989.	
PR	08-DEC-2000;	20000US-0251990.	
PR	11-DEC-2000;	20000US-0254097.	
PR	05-JAN-2001;	20010US-0259678.	
XX			
PA	(HUMA-)	HUMAN GENOME SCI INC.	
XX			
PI	Rosen CA,	Barash SC, Ruben SM;	
XX			
DR	WPI: 2001-488783/53.		
DR	P-PSDB; AAU16534.		
XX			
PT	New nucleic acid molecules encoding 461 human secreted proteins for		
PT	diagnosing, preventing, treating or ameliorating medical conditions and		
PT	used as food additives or preservatives -		
XX			
PS	Claim 1; SEQ ID NO 700; 980bp; English.		
XX			
CC	The invention relates to isolated nucleic acid molecules and their		
CC	encoded secreted proteins. The nucleic acids and proteins are used to		
CC	prevent, treat or ameliorate a medical condition in e.g. humans, mice,		
CC	rabbits, goats, horses, cats, dogs, chickens or sheep. They		
CC	are also used in diagnosing a pathological condition or susceptibility		
CC	to a pathological condition. Antibodies to the proteins can also		
CC	be used in alleviating symptoms associated with the disorders and in		
CC	diagnostic immunoassays e.g. radioimmunoassays or enzyme linked		
CC	immunoassay assays (ELISA). Disorders which are diagnosed or treated		
CC	include autoimmune diseases e.g. rheumatoid arthritis,		
CC	hyperproliferative disorders e.g. neoplasms of the breast or liver,		
CC	cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders		
CC	e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.		
CC	Alzheimer's disease, infections caused by bacteria, viruses and fungi		
CC	and ocular disorders e.g. corneal infection, and many other		
CC	disorders listed in the specification. The polypeptides can also		
CC	be used to aid wound healing and epithelial cell proliferation, to		
CC	prevent skin aging due to sunburn, to maintain organs before		
CC	transplantation, for supporting cell culture of primary tissues, to		
CC	regenerate tissues and in chemotaxis. The polypeptides can also be used		
CC	as a food additive or preservative to increase or decrease storage		
CC	capabilities, fat content, lipid, protein, carbohydrate, vitamins,		
CC	minerals, cofactors and other nutritional components. The present		
CC	sequence encodes a novel secreted protein of the invention.		
Query Match	0.8%;	Score 22;	DB 22; Length 758;
Best Local Similarity	100.0%;	Pred. No. 2.2;	
Matches 22; Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
QY	1419 cacacagagagagagcctta	g 1440	
Db	283 cacacagagagagagcctta	g 304	
RESULT 48			
ABA21456			
ID	ABA21456 standard; DNA; 1038 BP.		
XX			
AC	ABA21456;		
XX			
DT	23-JAN-2002 (first entry)		
XX			
DE	Human nervous system related polynucleotide SEQ ID NO 13787.		
XX			
KW	Human; nootropic; neuroprotective; cytosatic; dermatological; virocidic;		
KW	immunosuppressive; antikinflammatory; anti-HIV; antibacterial; antiviral;		
KW	antiparkinsonian; antischizofrenic; antianemic; antiallergic; cancer;		
KW	antirheumatic; hepatotropic; cerebroprotective; antinflammatory;		
KW	antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;		

KW anti-parasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX Homo sapiens.
XX WO200159063-A2.
XX
XX 16-AUG-2001.
XX
XX 17-JAN-2001; 2001MO-US01334.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0203515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232168.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0233240.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.

PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM.
XX
XX WPI: 2001-483426/52.
DR
XX
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
XX
PS Disclosure: SEQ ID NO 39080; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosstatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and

CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human Immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 1038 BP; 238 A; 242 C; 184 G; 374 T; 0 other;

Query Match 0.8%; Score 22; DB 22; Length 1038;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1251 cacactggggaggaagcctatg 1272
Db 843 CACACTGGGAGGAGCCTTATG 822

RESULT 50
AA571465
ID AA571465 standard; cDNA; 1185 BP.
XX
XX AA571465;
AC
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #7269.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
OS
XX
PN WO200175067-A2.
XX
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001MO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
DR P-PSDB; ABG07278.
XX
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX
PS Claim 1: SEQ ID NO 7269; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

50 Sequence 1185 BP; 344 A; 262 C; 295 G; 284 T; 0 other;

Query Match 0.8%; Score 22; DB 23; Length 1185;

Best Local Similarity 100.0%; Pred. No. 2.2;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1587 cacacgagggaagaccatttg 1608

Db 139 cacacgagggaagaccatttg 160

Search completed: May 16, 2002, 10:45:05
Job time: 8490 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2002, 08:20:35 ; Search time 2877.72 Seconds
(without alignments)
13001.112 Million cell updates/sec

Title: US-09-898-556A-3
Perfect score: 2772
Sequence: 1 cagcgcgcgttaagctgctg.....ttctaccaccctcaccct 2772

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 674847542 residues

Word size : 0
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

EST:*
1: em_estba:*
2: em_estbun:*
3: em_estin:*
4: em_estmu:*
5: em_estrov:*
6: em_estcpl:*
7: em_estcro:*
8: em_hic:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	721	26.0	1082	10	BM450709 AGENCOURT
2	677	24.4	1035	10	BM470740 AGENCOURT
3	585	21.1	1071	10	BE741389
4	581	21.0	590	9	AW245709
5	545	19.7	1009	10	BM449478 AGENCOURT
6	510	18.4	563	9	AI669533
7	509	18.4	535	9	AI6047439
8	492	17.7	558	9	AI802142
9	482	17.4	537	9	BE220142
10	471	17.0	624	9	BE178716
11	458	16.5	512	10	BE111754
12	454	16.4	505	10	BM463501
13	449	16.2	486	9	AW401692
14	419	15.1	486	9	AW720986
15	414	14.9	613	9	AW964787
16	413	14.9	697	10	BE391140
17	405	14.6	673	10	BC928720

C 18	402	14.5	453	9	AW513018	AW513018 xc76f03.x
C 19	399	14.4	481	9	AI276016	AI276016 qw08h02.x
C 20	395	14.2	451	9	AA450184	AA450184 zx42e09.x
C 21	383	13.8	435	9	AI081665	AI081665 ou63e05.x
C 22	383	13.8	449	9	A1750087	A1750087 alc5f07.x
C 23	378	13.6	455	10	BE501382	BE501382 hws1a02.x
C 24	378	13.6	728	10	BE746403	BE746403 601579525
C 25	372	13.4	424	9	AI016684	AI016684 ou96d05.x
C 26	372	13.4	458	9	AI440109	AI440109 t156h06.x
C 27	372	13.4	428	9	AI953623	AI953623 wg23a03.x
C 28	366	13.2	501	9	AW440124	AW440124 UI-H-BIL-
C 29	363	13.1	426	9	AI828662	AI828662 tr06f04.x
C 30	355	12.8	486	9	AI650983	AI650983 wa96d07.x
C 31	351	12.7	473	10	BE476173	BE476173 naa29c11
C 32	350	12.6	385	9	AA218658	AA218658 zq96g04.x
C 33	350	12.6	500	9	AI986444	AI986444 wz65d09.x
C 34	345	12.4	551	10	BE395838	BE395838 601310077
C 35	344	12.4	508	9	AI860707	AI860707 w115d12.x
C 36	343	12.4	399	9	AI289399	AI289399 qw32e05.x
C 37	340	12.3	392	9	AA218659	AA218659 zq96g04.x
C 38	340	12.3	395	9	AI361705	AI361705 qz18f04.x
C 39	337	12.2	513	9	AW958936	AW958936 EST371006
C 40	333	12.0	794	10	BI256434	BI256434 602974454
C 41	332	12.0	383	9	AA450118	AA450118 zx42e09.x
C 42	318	11.5	378	9	AA484745	AA484745 ne81h06.s
C 43	315	11.4	369	9	AI159662	AI159662 qb80f10.x
C 44	315	11.4	818	9	AL566746	AL566746 AL566746
C 45	312	11.3	501	10	R50617	R50617 yj57b01.r1
C 46	311	11.2	427	10	R83624	R83624 yp16h01.r1
C 47	308	11.1	410	9	AW513538	AW513538 x086b03.x
C 48	305	11.0	383	9	AW139497	AW139497 UI-H-BIL-
C 49	291	10.5	680	9	AW579734	AW579734 RC1-HT025
C 50	290	10.5	1041	10	BE798936	BE798936 601568319
C 51	279	10.1	712	10	BF026155	BF026155 601668851
C 52	279	10.1	761	11	BC010024	BC010024 Homo sapi
C 53	269	9.7	809	10	BE746802	BE746802 602704138
C 54	256	9.2	472	10	BF844319	BF844319 RC1-HT022
C 55	255	9.2	275	9	AW360766	AW360766 PM3-CT024
C 56	251	9.1	577	10	BE275642	BE275642 601121171
C 57	249	9.0	884	10	BG402277	BG402277 602465851
C 58	246	8.9	323	9	AI655251	AI655251 wb68a12.x
C 59	246	8.9	474	10	BF940868	BF940868 hu60b12.x
C 60	240	8.7	377	10	BE773364	BE773364 CM2-IT003
C 61	239	8.6	669	10	BE618307	BE618307 601462687
C 62	228	8.2	439	9	AW753060	AW753060 PM3-CT024
C 63	223	8.0	481	9	AW574555	AW574555 UI-HR-BKO
C 64	223	8.0	565	9	AA777271	AA777271 z196b07.s
C 65	215	7.8	451	10	BF768400	BF768400 QV3-IT000
C 66	209	7.5	317	9	AI917803	AI917803 wb53b08.x
C 67	209	7.5	456	9	AI393883	AI393883 t65h06.x
C 68	207	7.5	260	9	AI420536	AI420536 te99a06.x
C 69	205	7.4	530	9	AW954105	AW954105 EST366175
C 70	204	7.3	371	9	AI633088	AI633088 t201h05.x
C 71	202	7.3	371	9	AW868888	AW868888 MK1-SN006
C 72	201	7.3	300	9	AA384543	AA384543 EST98201
C 73	198	7.1	524	10	R90992	R90992 yp93b08.r1
C 74	198	7.1	267	9	AL537287	AL537287 AL537287
C 75	197	7.1	908	9	AI990117	AI990117 ws29d01.x
C 76	195	7.0	276	9	AW134770	AW134770 UI-H-BIL-
C 77	194	7.0	281	9	AA012980	AA012980 ze35h03.s
C 78	189	6.8	664	10	BE619323	BE619323 601473319
C 79	182	6.6	304	9	AI868995	AI868995 wb91h09.x
C 80	179	6.5	213	10	BC995316	BC995316 CM0-HT129
C 81	171	6.2	180	10	BI004708	BI004708 CM0-HN020
C 82	160	5.8	324	10	BF222885	BF222885 7q27a04.x
C 83	157	5.7	450	9	BE167882	BE167882 CM1-HT051
C 84	156	5.6	401	10	BC010675	BC010675 CM4-GN029
C 85	155	5.6	252	9	AI937720	AI937720 wp83b12.x
C 86	155	5.6	319	9	AW051156	AW051156 w208d02.x
C 87	152	5.5	326	10	BE352587	BE352587 IL3-HT061
C 88	149	5.4	200	10	BC995318	BC995318 CM0-HT129
C 89	147	5.3	260	10	BI006722	BI006722 CM3-RT000
C 90	146	5.3	599	10	BC470016	BC470016 602533494

91	136	4.9	163	9	AM179090	AM179090 MR4-ST006	164	23	0.8	300	9	AU099095
92	130	4.7	251	10	BE810472	BE810472 MR4-LT007	165	23	0.8	335	10	BF375372
93	128	4.6	336	10	BG945317	BG945317 PM0-AN008	166	23	0.8	435	9	AM128568
94	128	4.6	463	10	R88735	R88735 yp3p08.s1	167	23	0.8	436	10	T09047
95	124	4.5	246	9	AA013069	AA013069 ze35R03.r	168	23	0.8	446	9	AA584621
96	116	4.2	330	9	AA340342	AA340342 EST45533	169	23	0.8	519	9	AA244891
97	101	3.8	181	10	BF821542	BF821542 CM3-RT000	170	23	0.8	554	10	BE252919
98	106	3.6	489	10	BE258293	BE258293 601115970	171	23	0.8	569	9	A1332825
99	99	3.6	150	9	A1364896	A1364896 gx96h11.x	172	23	0.8	575	9	A1159431
100	97	3.5	158	9	AM594170	AM594170 h054a09.x	173	23	0.8	593	9	A1810272
101	96	3.5	428	10	RS0523	RS0523 yj57b01.s1	174	23	0.8	611	9	AA930937
102	90	3.2	637	10	BE005020	BE005020 MR0-BN011	175	23	0.8	627	12	BH342173
103	88	3.2	148	9	AM511484	AM511484 h089e10.x	176	23	0.8	650	9	AA161830
104	86	3.1	156	9	A1656184	A1656184 tc38e12.x	177	23	0.8	679	10	B1660479
105	78	2.8	337	9	AM243552	AM243552 2823003.3	178	23	0.8	698	9	AM281508
106	77	2.8	326	10	BF32587	BF32587 IL3-HT061	179	23	0.8	711	12	A2984765
107	68	2.5	126	9	AM887794	AM887794 PM2-CT009	180	23	0.8	724	10	BE264198
108	68	2.5	148	9	A1637511	A1637511 tc25g09.x	181	23	0.8	741	9	AM953464
109	58	2.1	174	9	AM007866	AM007866 wv46b05.x	182	23	0.8	817	10	B1454371
110	58	2.1	571	10	BG335907	BG335907 602404637	183	23	0.8	848	10	BG818388
111	53	1.9	622	10	BE964413	BE964413 601658119	184	23	0.8	850	10	BF346885
112	49	1.8	869	10	BG334472	BG334472 602461105	185	23	0.8	906	10	B1555319
113	46	1.7	996	10	BE963747	BE963747 601657087	186	22	0.8	173	10	BF959482
114	41	1.5	227	10	BG946561	BG946561 PM3-KT004	187	22	0.8	205	9	BB381517
115	37	1.3	159	10	BF876015	BF876015 IL3-ET011	188	22	0.8	229	9	AA495878
116	34	1.2	360	10	BF747325	BF747325 RC1-BT025	189	22	0.8	220	10	BG690002
117	33	1.2	395	9	AM866953	AM866953 CM3-SM003	190	22	0.8	245	10	AM955916
118	32	1.2	374	9	A1859123	A1859123 w133q08.x	191	22	0.8	289	9	AM955916
119	32	1.2	633	9	BC747151	BC747151 602704372	192	22	0.8	303	12	AQ042988
120	31	1.1	519	10	BM287346	BM287346 5277111.MA	193	22	0.8	318	10	BE884570
121	29	1.0	276	9	AA385857	AA385857 EST99634	194	22	0.8	323	12	BH274131
122	29	1.0	295	10	T23204	T23204 C2H2-260T.H	195	22	0.8	339	12	AQ061946
123	29	1.0	692	10	B1918302	B1918302 603182986	196	22	0.8	383	10	BF433559
124	29	1.0	810	10	BG697869	BG697869 602661085	197	22	0.8	398	10	BG897513
125	29	1.0	860	10	B1767915	B1767915 603061059	198	22	0.8	409	9	AA457311
126	29	1.0	910	10	B1669973	B1669973 603394125	199	22	0.8	411	10	F18480
127	29	1.0	975	10	BF345414	BF345414 602018819	200	22	0.8	416	12	AQ415870
128	28	1.0	325	10	BF806376	BF806376 RC2-CI008	201	22	0.8	428	12	AC089708
129	28	1.0	440	9	AV659318	AV659318 AV659318	202	22	0.8	442	10	BF741901
130	28	1.0	459	12	AQ184830	AQ184830 HS.2216.B	203	22	0.8	446	10	BF858002
131	28	1.0	461	10	BE255957	BE255957 601113340	204	22	0.8	451	10	BE709979
132	28	1.0	520	10	BE250926	BE250926 601107374	205	22	0.8	453	9	AA632636
133	28	1.0	569	10	BE251006	BE251006 601107411	206	22	0.8	454	12	AQ466027
134	28	1.0	614	10	BE257108	BE257108 601108841	207	22	0.8	456	9	AM073513
135	28	1.0	711	12	AG090673	AG090673 Pan.trog1	208	22	0.8	457	12	AO624861
136	28	1.0	817	10	BE744885	BE744885 601576334	209	22	0.8	481	10	BF8990400
137	28	1.0	919	10	BG396806	BG396806 602459607	210	22	0.8	487	9	AA983513
138	28	1.0	1028	10	BG329013	BG329013 602428674	211	22	0.8	487	12	AQ207966
139	27	1.0	189	9	AW751834	AW751834 QV1-CT011	212	22	0.8	494	9	A1014809
140	27	1.0	266	9	AL449531	AL449531 AL449531	213	22	0.8	505	10	BE858098
141	27	1.0	265	9	AL449532	AL449532 AL449532	214	22	0.8	506	12	A2519608
142	27	1.0	618	9	AW751839	AW751839 QV1-CT011	215	22	0.8	510	9	A1636158
143	27	1.0	612	10	BE536073	BE536073 601061719	216	22	0.8	519	9	A1636158
144	27	1.0	782	10	BG681219	BG681219 602850828	217	22	0.8	523	9	AA476780
145	27	1.0	844	10	BI087151	BI087151 602850828	218	22	0.8	534	9	A1263449
146	26	0.9	474	10	BG472668	BG472668 602514470	219	22	0.8	545	12	AQ054920
147	25	0.9	439	10	BI340197	BI340197 365393.MA	220	22	0.8	549	10	BF437602
148	25	0.9	476	10	T70255	T70255 yd46b03.s1	221	22	0.8	549	12	AQ772417
149	25	0.9	539	10	BI340196	BI340196 365393.MA	222	22	0.8	550	10	BE304164
150	25	0.9	561	9	A1668883	A1668883 wc75b09.x	223	22	0.8	559	10	B1540243
151	25	0.9	606	10	BF766794	BF766794 RC5-CS002	224	22	0.8	573	10	BE927220
152	25	0.9	613	10	BE250978	BE250978 601107057	225	22	0.8	574	12	AO697253
153	25	0.9	651	12	AG079437	AG079437 Pan.trog1	226	22	0.8	583	12	AZ080864
154	25	0.9	702	12	AG039049	AG039049 Pan.trog1	227	22	0.8	588	10	W22252
155	25	0.9	911	10	BE513012	BE513012 601172429	228	22	0.8	594	10	BG688499
156	25	0.9	911	10	BE513012	BE513012 601172429	229	22	0.8	612	12	AQ310487
157	24	0.9	427	10	T81042	T81042 yd26b03.r1	230	22	0.8	622	12	AQ055307
158	24	0.9	450	10	T84260	T84260 yd31909.t1	231	22	0.8	626	12	AZ288125
159	24	0.9	627	12	AG0594180	AG0594180 HS.2095.A	232	22	0.8	636	12	AQ394703
160	24	0.9	638	12	AG033590	AG033590 Pan.trog1	233	22	0.8	657	12	BI159507
161	24	0.9	650	12	AG059861	AG059861 Pan.trog1	234	22	0.8	659	12	BH340309
162	23	0.8	229	9	AA341469	AA341469 EST46876	235	22	0.8	661	12	AG094858
163	23	0.8	237	12	AQ474928	AQ474928 CITR1-EL-	236	22	0.8			

C 383	20	0.7	410	9	AV837142	AV837142	AV837142	20	0.7	620	12	CNS02DR5	AL192794	Tetraodon
C 384	20	0.7	415	10	BF812712	BF812712	BF812712	20	0.7	625	10	BF399499	BF399499	UI-R-CAL-
C 385	20	0.7	416	10	BG554285	BG554285	BG554285	20	0.7	628	12	BF723386	BF723386	CH230-162
C 386	20	0.7	419	12	AO669911	AO669911	AO669911	20	0.7	629	9	AW078416	AW078416	f449c11.Y
C 387	20	0.7	426	12	AO346128	AO346128	AO346128	20	0.7	631	12	AZ442099	AZ442099	1M0234607
C 388	20	0.7	429	10	W86428	W86428	W86428	20	0.7	637	10	BG026033	BG026033	602292216
C 389	20	0.7	431	9	AA612532	AA612532	AA612532	20	0.7	641	10	BF683419	BF683419	602193587
C 390	20	0.7	439	9	A1548657	A1548657	A1548657	20	0.7	643	12	AZ361725	AZ361725	1M0106701
C 391	20	0.7	439	12	BH362923	BH362923	BH362923	20	0.7	644	12	AG128005	AG128005	Pan trogl
C 392	20	0.7	440	9	AW250546	AW250546	AW250546	20	0.7	650	12	AZ042772	AZ042772	RPCI-23-2
C 393	20	0.7	442	10	BM254749	BM254749	BM254749	20	0.7	656	12	AZ999962	AZ999962	2M0287105
C 394	20	0.7	447	10	T88890	T88890	T88890	20	0.7	658	9	AW962055	AW962055	EST374128
C 395	20	0.7	454	9	A1603959	A1603959	A1603959	20	0.7	663	10	BG429615	BG429615	602501301
C 396	20	0.7	455	9	AW197626	AW197626	AW197626	20	0.7	668	9	BB618073	BB618073	BB618073
C 397	20	0.7	458	10	BM147779	BM147779	BM147779	20	0.7	670	9	BB067752	BB067752	BB067752
C 398	20	0.7	462	10	H40464	H40464	H40464	20	0.7	674	10	AZ204137	AZ204137	SP 0097_B
C 399	20	0.7	464	9	A1715190	A1715190	A1715190	20	0.7	677	9	AL655196	AL655196	AL655196
C 400	20	0.7	467	9	AW206752	AW206752	AW206752	20	0.7	681	10	BE408214	BE408214	601302537
C 401	20	0.7	468	10	B1883489	B1883489	B1883489	20	0.7	688	10	BF294090	BF294090	002PBD05
C 402	20	0.7	486	10	BF045373	BF045373	BF045373	20	0.7	701	10	BG257604	BG257604	602377283
C 403	20	0.7	490	9	AW500608	AW500608	AW500608	20	0.7	717	9	BB639365	BB639365	BB639365
C 404	20	0.7	493	9	AM165414	AM165414	AM165414	20	0.7	721	10	B1916631	B1916631	603178569
C 405	20	0.7	497	10	BG653680	BG653680	BG653680	20	0.7	723	10	B1067915	B1067915	B1067915
C 406	20	0.7	497	10	BM364215	BM364215	BM364215	20	0.7	726	10	BF723490	BF723490	mab32f05
C 407	20	0.7	501	9	A1883736	A1883736	A1883736	20	0.7	729	10	BG393583	BG393583	602411979
C 408	20	0.7	501	10	H30708	H30708	H30708	20	0.7	731	10	B1739075	B1739075	603361335
C 409	20	0.7	506	9	AA852047	AA852047	AA852047	20	0.7	751	12	BH329854	BH329854	CH230-124
C 410	20	0.7	507	9	A1930163	A1930163	A1930163	20	0.7	755	10	B1889250	B1889250	2F637-2-0
C 411	20	0.7	511	10	BF015449	BF015449	BF015449	20	0.7	761	9	AL537987	AL537987	AL537987
C 412	20	0.7	515	9	A1092527	A1092527	A1092527	20	0.7	774	10	BM046555	BM046555	603627188
C 413	20	0.7	518	9	A1188459	A1188459	A1188459	20	0.7	778	10	BG771829	BG771829	CH2270511
C 414	20	0.7	519	9	AW146179	AW146179	AW146179	20	0.7	781	12	BH271885	BH271885	60230-107
C 415	20	0.7	519	9	AM304314	AM304314	AM304314	20	0.7	789	10	BG109145	BG109145	602281341
C 416	20	0.7	521	10	B1062348	B1062348	B1062348	20	0.7	799	10	BG782256	BG782256	SEAWUC002
C 417	20	0.7	523	10	BM225410	BM225410	BM225410	20	0.7	800	10	B1767963	B1767963	603061334
C 418	20	0.7	528	12	AO552360	AO552360	AO552360	20	0.7	803	10	BE572656	BE572656	601329433
C 419	20	0.7	529	9	AM302887	AM302887	AM302887	20	0.7	804	12	CNS072EQ	CNS072EQ	603086646
C 420	20	0.7	532	12	AO604617	AO604617	AO604617	20	0.7	819	10	B1838075	B1838075	AL568497
C 421	20	0.7	536	10	BM089434	BM089434	BM089434	20	0.7	821	9	AL568497	AL568497	AL568497
C 422	20	0.7	537	10	BM142454	BM142454	BM142454	20	0.7	827	10	BF972810	BF972810	602241080
C 423	20	0.7	538	10	B1343192	B1343192	B1343192	20	0.7	837	10	B1150768	B1150768	602914858
C 424	20	0.7	539	10	BM273186	BM273186	BM273186	20	0.7	837	10	BE408269	BE408269	601302628
C 425	20	0.7	539	12	AZ294090	AZ294090	AZ294090	20	0.7	837	10	B1830259	B1830259	603072977
C 426	20	0.7	540	10	B1064152	B1064152	B1064152	20	0.7	840	10	BG175180	BG175180	602337502
C 427	20	0.7	541	9	AA402531	AA402531	AA402531	20	0.7	842	10	BE313621	BE313621	601153261
C 428	20	0.7	543	9	AM654426	AM654426	AM654426	20	0.7	855	10	BG702063	BG702063	602683879
C 429	20	0.7	550	10	BG086136	BG086136	BG086136	20	0.7	865	10	BF182646	BF182646	601809420
C 430	20	0.7	551	12	AZ955502	AZ955502	AZ955502	20	0.7	880	10	BF972892	BF972892	602241180
C 431	20	0.7	554	9	AW751840	AW751840	AW751840	20	0.7	889	10	BG622387	BG622387	602647079
C 432	20	0.7	555	9	AU211253	AU211253	AU211253	20	0.7	892	10	B1020016	B1020016	B1020016
C 433	20	0.7	556	9	AM652514	AM652514	AM652514	20	0.7	898	12	CNS02CLO	CNS02CLO	AL191277
C 434	20	0.7	561	9	BB639388	BB639388	BB639388	20	0.7	900	10	BG490940	BG490940	602519744
C 435	20	0.7	562	12	AZ040916	AZ040916	AZ040916	20	0.7	901	10	BE729270	BE729270	6021561627
C 436	20	0.7	571	9	BE141059	BE141059	BE141059	20	0.7	913	10	BF6694306	BF6694306	602082946
C 437	20	0.7	571	12	BM265301	BM265301	BM265301	20	0.7	927	9	AL524526	AL524526	AL524526
C 438	20	0.7	572	10	BF180853	BF180853	BF180853	20	0.7	927	12	CNS02M95	CNS02M95	AL203810
C 439	20	0.7	576	10	BE783521	BE783521	BE783521	20	0.7	930	10	BF183055	BF183055	601809627
C 440	20	0.7	576	12	AZ439533	AZ439533	AZ439533	20	0.7	936	10	B1656760	B1656760	603281431
C 441	20	0.7	582	12	BH047089	BH047089	BH047089	20	0.7	940	10	BF107705	BF107705	601823668
C 442	20	0.7	583	10	BM494784	BM494784	BM494784	20	0.7	940	10	BF683074	BF683074	602139188
C 443	20	0.7	587	9	AL119330	AL119330	AL119330	20	0.7	951	10	BF966463	BF966463	602287084
C 444	20	0.7	588	10	BE816540	BE816540	BE816540	20	0.7	972	11	BC013605	BC013605	602095446
C 445	20	0.7	600	9	AL120716	AL120716	AL120716	20	0.7	974	10	BF578885	BF578885	602095446
C 446	20	0.7	602	9	BG805595	BG805595	BG805595	20	0.7	999	10	BG335049	BG335049	602403506
C 447	20	0.7	602	9	AA772083	AA772083	AA772083	20	0.7	1059	12	CNS073RK	CNS073RK	AL27878
C 448	20	0.7	605	10	BG675577	BG675577	BG675577	20	0.7	1062	10	BM451528	BM451528	clone BAO
C 449	20	0.7	614	9	AM764673	AM764673	AM764673	20	0.7	1062	9	AL524527	AL524527	AL524527
C 450	20	0.7	615	10	BM049184	BM049184	BM049184	20	0.7	1113	10	BF684931	BF684931	602142066
C 451	20	0.7	615	10	BE314293	BE314293	BE314293	20	0.7	1206	10	BG245517	BG245517	602355076
C 452	20	0.7	618	10	BM049625	BM049625	BM049625	20	0.7	1321	10	BM473019	BM473019	AGENCOURT
C 453	20	0.7	620	9	BB655489	BB655489	BB655489	20	0.7	1431	9	A1964256	A1964256	EST263370
C 454	20	0.7	620	10	BG087846	BG087846	BG087846	20	0.7	1556	20	BE966977	BE966977	60160708
C 455	20	0.7	620	10	B1890167	B1890167	B1890167	20	0.7	1769	11	AK016063	AK016063	Mus muscu

821	19	0.7	728	10	BE369136	BE369136 601222926	894	19	0.7	1179	10	BM479116	BM479116
822	19	0.7	731	9	BB574195	BB574195	895	19	0.7	1188	10	BG166903	BG166903
823	19	0.7	731	12	AG109464	AG109464 Pan trogl	896	19	0.7	1515	11	AK018294	AK018294
824	19	0.7	735	12	BH285208	BH285208 CH230-104	897	19	0.7	1843	10	BF128533	BF128533
825	19	0.7	739	9	BB328667	BB328667	898	19	0.7	2048	11	BC017051	BC017051
826	19	0.7	741	10	BE276357	BE276357 B03144005	899	19	0.6	118	10	BE939421	BE939421
827	19	0.7	741	12	BH566840	BH566840 BOGMM47TR	900	18	0.6	123	12	AZ638864	AZ638864
828	19	0.7	745	12	AZ324806	AZ324806 IM0046120	901	18	0.6	132	10	BF090480	BF090480
829	19	0.7	748	10	BG832225	BG832225 602765605	902	18	0.6	141	10	BG007755	BG007755
830	19	0.7	750	12	AZ731909	AZ731909 RPT-24-1	903	18	0.6	141	12	BH010314	BH010314
831	19	0.7	753	9	AU143397	AU143397	904	18	0.6	135	9	AA371471	AA371471
832	19	0.7	754	10	BI970606	BI970606 GM830011A	905	18	0.6	161	9	AM561898	AM561898
833	19	0.7	762	10	BG419708	BG419708 602451939	906	18	0.6	162	10	BF587323	BF587323
834	19	0.7	768	10	BG536481	BG536481 602564582	907	18	0.6	164	9	BA088845	BA088845
835	19	0.7	768	10	BG177764	BG177764 602698053	908	18	0.6	166	9	AI910059	AI910059
836	19	0.7	773	10	BI259848	BI259848 602971507	909	18	0.6	167	12	AZ287256	AZ287256
837	19	0.7	775	10	BI489574	BI489574 603032008	910	18	0.6	169	10	BI022405	BI022405
838	19	0.7	776	10	BI968622	BI968622 GM830005B	911	18	0.6	174	10	BE768696	BE768696
839	19	0.7	787	10	BE887288	BE887288 601510111	912	18	0.6	175	10	BI708541	BI708541
840	19	0.7	791	12	AQ741109	AQ741109 HS-2272-A	913	18	0.6	175	10	BI709210	BI709210
841	19	0.7	792	9	AU130601	AU130601	914	18	0.6	176	10	BI840229	BI840229
842	19	0.7	800	10	BE794809	BE794809 601589269	915	18	0.6	178	9	AA284530	AA284530
843	19	0.7	807	10	BI227029	BI227029 602952184	916	18	0.6	181	9	AA213773	AA213773
844	19	0.7	808	10	BG777192	BG777192 602664414	917	18	0.6	181	12	AZ563544	AZ563544
845	19	0.7	817	10	BF791220	BF791220 602250735	918	18	0.6	195	10	BG218342	BG218342
846	19	0.7	818	9	AL048542	AL048542 DKF2P5860	919	18	0.6	195	12	AZ803588	AZ803588
847	19	0.7	818	10	BI522164	BI522164 603081575	920	18	0.6	197	9	BA256625	BA256625
848	19	0.7	830	10	BG432658	BG432658 602496329	921	18	0.6	200	9	AA333524	AA333524
849	19	0.7	838	10	BI408930	BI408930 602966412	922	18	0.6	202	10	BG202154	BG202154
850	19	0.7	839	11	AK007385	AK007385 Mus muscu	923	18	0.6	203	10	BF411599	BF411599
851	19	0.7	840	9	AU132370	AU132370	924	18	0.6	204	9	AA616935	AA616935
852	19	0.7	845	9	AU130169	AU130169	925	18	0.6	206	10	BG212632	BG212632
853	19	0.7	849	10	BI838890	BI838890 603087272	926	18	0.6	207	10	BF736582	BF736582
854	19	0.7	850	12	AO900025	AO900025 HS-3162-A	927	18	0.6	215	9	AW026744	AW026744
855	19	0.7	853	10	BG924061	BG924061 602822278	928	18	0.6	217	9	BB038977	BB038977
856	19	0.7	859	10	BF673336	BF673336 602136194	929	18	0.6	219	9	AI915039	AI915039
857	19	0.7	862	10	BG683381	BG683381 602651482	930	18	0.6	220	10	BE303653	BE303653
858	19	0.7	866	10	BE253154	BE253154 601116687	931	18	0.6	223	10	BF744093	BF744093
859	19	0.7	868	10	BF684872	BF684872 602142589	932	18	0.6	224	12	AZ222734	AZ222734
860	19	0.7	870	12	CNS03TYO	AL260453 Tetradon	933	18	0.6	227	9	AI962077	AI962077
861	19	0.7	872	9	AL539493	AL539493	934	18	0.6	228	10	BG006987	BG006987
862	19	0.7	872	10	BG772120	BG772120 602721568	935	18	0.6	228	12	AQ342298	AQ342298
863	19	0.7	898	10	BF969103	BF969103 602270057	936	18	0.6	231	12	AQ026938	AQ026938
864	19	0.7	899	10	BI456360	BI456360 603172870	937	18	0.6	233	10	TQ0346	TQ0346
865	19	0.7	903	10	BG117671	BG117671 602349630	938	18	0.6	234	9	BB554958	BB554958
866	19	0.7	905	10	BI729498	BI729498 603349142	939	18	0.6	235	10	D20836	D20836
867	19	0.7	922	9	AU132930	AU132930	940	18	0.6	241	9	AV281527	AV281527
868	19	0.7	931	12	CNS06DDT	ALJ39592 T3 end of	941	18	0.6	241	9	BB084006	BB084006
869	19	0.7	932	10	BG259881	BG259881 602371838	942	18	0.6	244	10	TQ0318	TQ0318
870	19	0.7	934	10	BF684705	BF684705 602141081	943	18	0.6	244	12	AZ644493	AZ644493
871	19	0.7	945	9	BE973483	BE973483 601652319	944	18	0.6	246	9	AM590258	AM590258
872	19	0.7	947	9	AU067250	AU067250	945	18	0.6	247	9	AA844604	AA844604
873	19	0.7	949	10	BF666815	BF666815 602140781	946	18	0.6	248	9	AI637899	AI637899
874	19	0.7	952	9	BB618414	BB618414	947	18	0.6	252	10	F34684	F34684
875	19	0.7	959	9	AU167354	AU167354	948	18	0.6	253	9	AA076896	AA076896
876	19	0.7	960	10	BM475825	BM475825	949	18	0.6	255	10	BM042601	BM042601
877	19	0.7	972	10	BF976622	BF976622 602244114	950	18	0.6	257	9	AV061040	AV061040
878	19	0.7	975	10	BG115887	BG115887 602317303	951	18	0.6	258	10	BE672134	BE672134
879	19	0.7	979	10	BG576433	BG576433 602595675	952	18	0.6	259	9	AI825519	AI825519
880	19	0.7	980	10	BG117509	BG117509 602347637	953	18	0.6	263	9	AI019402	AI019402
881	19	0.7	991	10	BG242226	BG242226 602334590	954	18	0.6	263	9	AI623173	AI623173
882	19	0.7	991	10	BG683980	BG683980 602651482	955	18	0.6	264	10	BG218979	BG218979
883	19	0.7	992	10	BE893396	BE893396 601436727	956	18	0.6	265	9	AA078649	AA078649
884	19	0.7	996	12	CNS012RR	AL101937 Drosophila	957	18	0.6	265	9	BB004173	BB004173
885	19	0.7	997	10	BG758928	BG758928 602713040	958	18	0.6	265	9	BB382159	BB382159
886	19	0.7	1012	12	BI054588	BI054588 ENTVO4T	959	18	0.6	269	9	AM196911	AM196911
887	19	0.7	1012	10	BI090312	BI090312 602855419	960	18	0.6	269	10	T27722	T27722
888	19	0.7	1020	10	BI759069	BI759069 603042864	961	18	0.6	271	9	BB441707	BB441707
889	19	0.7	1023	10	BG342230	BG342230 602374307	962	18	0.6	272	9	AI956496	AI956496
890	19	0.7	1072	10	BF339506	BF339506 602039163	963	18	0.6	273	9	BB473487	BB473487
891	19	0.7	1084	10	BE562821	BE562821 601336529	964	18	0.6	274	10	BG187038	BG187038
892	19	0.7	1095	10	BE791043	BE791043 601586504	965	18	0.6	274	10	BF998461	BF998461
893	19	0.7	1101	12	CNS05JIT	ALJ340238 Tetradon	966	18	0.6	277	9	BB089729	BB089729

967	18	0.6	278	9	AM304018	xv15h11.x
c 968	18	0.6	278	9	BB564825	BB564825
969	18	0.6	278	10	R10355	VF32412.r1
c 970	18	0.6	279	9	AV260992	AV260992
971	18	0.6	279	10	BG211877	BG211877
c 972	18	0.6	280	10	BG193003	RST12128
973	18	0.6	281	9	AV144913	AV144913
c 974	18	0.6	281	9	BA154476	BA154476
c 975	18	0.6	283	9	AI800894	wq14a12.x
c 976	18	0.6	284	9	AA843970	a191a02.s
c 977	18	0.6	284	9	AA312592	AA312592
c 978	18	0.6	284	10	F03657	HSC25D102.n
c 979	18	0.6	284	12	A2293429	PCPI-23-1
c 980	18	0.6	284	12	A2575767	AST-T32B0
981	18	0.6	285	9	BB360814	BB360814
c 982	18	0.6	285	10	BE571374	601330205
983	18	0.6	286	10	BE768710	QV2-FT001
984	18	0.6	288	9	AW003468	wq66b08.x
985	18	0.6	288	10	BE502516	h219f09.x
986	18	0.6	289	9	AI652715	wb30h08.x
987	18	0.6	289	9	AV122743	AV122743
988	18	0.6	290	10	BG203005	RST22374
c 989	18	0.6	291	9	AA334780	EST39041
c 990	18	0.6	292	9	AW669561	MR3-SN006
c 991	18	0.6	293	10	BG208440	RST27941
992	18	0.6	296	10	BF931762	CM2-NF021
c 993	18	0.6	298	9	AI475015	tm62b10.x
994	18	0.6	298	10	N75923	y264e10.r1
995	18	0.6	299	10	Z21091	HSAAADML0.T
c 996	18	0.6	300	9	AI244521	qk14c09.x
997	18	0.6	301	9	BB391078	BB391078
998	18	0.6	302	9	AA833475	uc91b11.r
c 999	18	0.6	302	9	BB116648	BB116648
1000	18	0.6	302	9	BB519406	BB519406

ALIGNMENTS

RESULT 1	BM450709	1082 bp	mRNA	linear	EST 05-FEB-2002
LOCUS	BM450709	1082 bp	mRNA	linear	EST 05-FEB-2002
DEFINITION	AGENCOURT_6394734 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5494415				
ACCESSION	BM450709				
VERSION	BM450709.1	GI:18499749			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
TITLE	1 (bases 1 to 1082)				
JOURNAL	NIH-MGC http://mgi.nci.nih.gov/				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)				
	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: csapbs-remail.nih.gov				
	Tissue Procurement: ATCC				
	cDNA Library Preparation: Life Technologies, Inc.				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)				
	DNA Sequencing by: Agencourt Bioscience Corporation				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LLNL at:				
	http://image.llnl.gov				
	Plate: L1AM12119 row: k column: 24				
	High quality sequence stop: 585.				
FEATURES	Location/Qualifiers				
SOURCE	1..1082				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone_image="5494415"				
	/clone_lib="NIH_MGC_67"				
	/tissue_type="retinoblastoma"				

/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

BASE COUNT 298 a 261 c 292 g 226 t 5 others

ORIGIN

Query Match 25.0%; Score 721; DB 10; Length 1082;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 721; Conservative 0;

Qy	849	atcaaatatgaagatttgccagcgttatacaaggagcaaacctcttagctccag	908		
Db	2	ATCAAAATATGAAGAGCTTTGGCCAGCTTTATCAAGAGATCAAACTCTTAGCTCCAG	61		
Qy	909	aagacacaaactgggagacacttacaatgtaactgagtgaggagacagcttggcagt	968		
Db	62	AAGACACAAACTGGGAGACACTTACATGTACATGAGTGGGAGACAGCTTTGGCAGT	121		
Qy	969	atgtcagtcctcatacaaaacccaaggagacacactctgggggaaagccttatgtgtcagg	1028		
Db	122	ATGTCAAGTCTTCATCAAAAACCAAGACACACTCTGGGGAAACCTTATGTGTGCAAG	181		
Qy	1029	gaatgtggcggcagcgttactcgttggaatcaaacctgatacacaatcagaagacacacta	1088		
Db	182	GAAATGTGGCGGAGGCTTTATGCTGAAGTCAAACTGATACATCAGACAGACACATCA	241		
Qy	1089	ggggagaaaccttatgtgtgcaaggattgtgagcagagccttacttggaaagtcgaacctc	1148		
Db	242	GGGGAGAAACCTTATGTGTGCAAGATTTGTGACGAGCGCTTATGTGAAGTCGAACCTC	301		
Qy	1149	tttaacatcagcggacacacactcaaggcctttagtgttgaaagaaatgtgtggcag	1208		
Db	302	TTTACACATCAGCGGACACACTGAGGCTCAAGCCTTATGTGTGAAGAAATGTGGCAG	361		
Qy	1209	agcctttagccttaagtcacaaccttataccacacagagggcgacacacttggagagaacct	1268		
Db	362	AGCTTAGCCTTACATCAAACTCAATTAACCAACAGAGGCGCCACACTGGAGAAAGCCT	421		
Qy	1269	tatgttttcaggaagtgtgtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	1328		
Db	422	TATGTTTTGGAGGGAATGTGGGGGCTTCCCGCATTCACACTGTGTCAGACACAAG	481		
Qy	1329	agggacacatcagagagaagagccttaccatttgcaggaggtgtgagcaaggccttagccag	1388		
Db	482	AGGACACATTCAGAGAGAGAGCCTTACATTTGCAAGGAGGTGTGACCAAGCCTTTAGCCAG	541		
Qy	1389	aagtcacacatcatalcagaacacttaaggacacacacagagagaagaagccttatgtatgcaca	1448		
Db	542	AAGTCACACCTCATCAGACACTTAAGAGACACACAGAGAGAAACCTTATGTATGACACA	601		
Qy	1449	gaatgtggcggtcactttagctgtgaaatcaaacctcaaaaacaccacagagagacactca	1508		
Db	602	GAAATGTGGCGGCTTATGTGTGCTGGAATCAAACTCAAAACACACACAGGACACACTCA	661		
Qy	1509	gggggttaaaccttatgtctgccttgagtgagtgagtgagtgagtgagtgagtgagtgagtg	1568		
Db	662	GGGGTTAAACCTTATGTGTGCTGCTGAGGTGGGGGAGTGTGCTTATGCTTGAAGTCAAACTT	721		
Qy	1569	a 1569			
Db	722	A 722			

RESULT 2	BM470740	1035 bp	mRNA	linear	EST 05-FEB-2002
LOCUS	BM470740	1035 bp	mRNA	linear	EST 05-FEB-2002
DEFINITION	AGENCOURT_6475358 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5579278				
ACCESSION	BM470740				
VERSION	BM470740.1	GI:18519782			

KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 1035)
TITLE	NIH-MGC http://mgc.nci.nih.gov/
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: csapbs-r@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: LLAM12335 row: k column: 23 High quality sequence stop: 653.
FEATURES	
source	1..1035 Location/Qualifiers /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5579278" /clone_lib="NIH-MGC-67" /tissue_type="retinoblastoma" /lab_host="DH10B (phage-resistant)" /note="Organ: eye; Vector: PCMV-SPOrt6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."
BASE COUNT	259 a 249 c 268 g 255 t 4 others
ORIGIN	

Query Match	24.4%	Score 677	DB 10	Length 1035
Best Local Similarity	100.0%	Pred. No. 0		
Matches 677	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 1655	cacgacccagaagacacactcaggggagaagacattgtagtgcgtgagtgfgyacgaag	1714		
DB 25	CACGCACCCAGAGACACACTCAGGGGAGAGACCATTTGTATGTGCTGAGTGTGGACGAG	84		
QY 1715	cttaatgataagtcacacccatcattcacaccagaagacacatcaagggaaaagccctt	1774		
DB 85	CTTTATGATGATGATGCACCCCTCATTTTCACACAGAGACACATTCAGGGAGAAAGCCTT	144		
QY 1775	tatgtgcagggagtgtytgcaagaagtttcgcgaagaagcctaaccgtttaggcacaagag	1834		
DB 145	TATGTGCAGAGGAGTGTGGCAGAAAGTTTCGCAGAAAGCCTTAACCTGTATAGGCACAAAG	204		
QY 1835	ggcacactaagtgcccttgtytgcaaggagtgtytgccaagcccttltgfyctaagttaac	1894		
DB 205	GGCACACTAGGTGCTTTGTGTGCAGGAGTGTGGCAAGCCTTTTGTGCTAAGTTAAC	264		
QY 1895	tctcatctaaacacagaagacacagcagggggagacccctcagtgtgcaggagtgtyg	1954		
DB 265	TCTCATTTAAACACACAGAGACACAGCAGGGGGAAAGCCTCATGTGTGTCAGGGAGTGTG	324		
QY 1955	gcaaggtcttagcggcagtcacacctcatatagacaccagaagacacatlcaggagaag	2014		
DB 325	GCAAGGCTTTAGCCGGCAGTCTACACCTCTATTGAGACACAGAGACACATTCAGGAGAGA	384		
QY 2015	gacctatatttcgaaaagtgfygagcggggcctttagtcggaagtcacaactcatagaca	2074		
DB 385	GCCTTATATTTCACAGAAATGTGGACGGGGCTTTAGTCGAGATCCAACTTATACAGAA	444		
QY 2075	tcaggagacacactcagagatagaactatgtatgatataggaatfygtfacaagccctttag	2134		
DB 445	TCAGAGGACACACTCAGGATAGAAACTTATGTGTATACGGAATGTGTACAGCCTTAGG	504		
QY 2135	ccaggagtcatactcatcagaaccagaagacacacacagtgctgtgtgcttlttcagcc	2194		

Db	505	CCAGAGAGTCATCTCTCATTCACAGACACCAAGACACACACAGTCGCTGTGGCTTTTTCACGC	564
QY	2195	attgtctaatataccaaagtgtgagacatctctgtgtgtgtatcatgcatgagacgttactgtt	2254
Db	565	ATTGCTAGATACCAAGAAGTGAGACATCTCGTGTGTGATATATGACATGAGACTGTACTGTGA	624
QY	2255	agacttgcattctccatccacccgtgagagaaatgtgtctattttgagagacccctggcc	2314
Db	625	ACACTGTGATCTCCATCCACCGTGAAGAGAAATTGCTGCTCATTTTTCAGAGACCCCTGGCC	684
QY	2315	ttccctcactgtgtgacgtg	2331
Db	685	TTCCCTCAGCTGTGATGG	701
RESULT	3		
BE741389			
LOCUS	BE741389	1071 bp	mrna
DEFINITION	601594255F1 NIH_MGC_9 Homo sapiens CDNA clone IMAGE:3948468 5',	linear	EST 15-SEP-2000
ACCESSION	BE741389		
VERSION	BE741389.1	GI:10155381	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 1071)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	unpublished (1998)		
	Contact: Robert Strausberg, Ph.D.		

Qy	1738	tttcaaccgaaggagacattcaagggaataacccctttatgtcgcggagagtgtgcagaa	1797
Db	36	GGGAGAGCACTTGTGATGTGCTGAGTGGAGCGAGGCTTTAATGATAAGTCCACCCITCA	95
Qy	1678	gggagaagccattgcatgctgtgagtgctgacgagagcttaagtataagtcaccctca	1737
Db	36	GGGAGAGCACTTGTGATGTGCTGAGTGGAGCGAGGCTTTAATGATAAGTCCACCCITCA	95
Qy	1738	tttcaaccgaaggagacattcaagggaataacccctttatgtcgcggagagtgtgcagaa	1797
Db	96	TTTTTCAACCCAGGAGGACACTTTCAGGGGGAATAACCCCTTTTATATGTGCAGAGGAGTGTGCACAGA	155

```
QY 1798 ggttcgcgaagaaccttaacctgtttaggcacagaaggacacactcaagtgctctgtc 1857
      |||||||
Db 156 ggTTTGGGAGAGAGCTTAACCTTTTAGGACACAGAGGACACTCAGGCTTTGTGT 215
QY 1858 gcaaggagtgctggcgaaggctttgtcttaagtaagtaacctcaataacacagaagcac 1917
      |||||||
Db 216 GCAGGAGAGTGTGGCAAGGCTTTGTGTAGTAAGTAACCTTCATTAAACACAGAGAGAC 275
QY 1918 aagcagaagggaagcctcatgtgtcgaaggagtggtggcgaaggctttagccggcaatcac 1977
      |||||||
Db 276 ACACGAGGGGGAGAGCTTCATGTGTGAGGAGAGTGCGGCAAGGCTTTAGCCGGCACTCAC 335
QY 1978 accctcatagacacacagagacacatcagagaagaagccttatcttgagaagaatgtc 2037
      |||||||
Db 336 ACCTCATTTAGACACACAGAGACATTCACAGAGAGAAAGCTTTATTTCAGAGAAAGTGTC 395
QY 2038 gacggggctttagtcggaagtcacacctatcacacatcagaagaacactcagataga 2097
      |||||||
Db 396 GACGGGGCTTTAGTCCGAAGTCCAACTTATCAGACATCCACAGACACACTCAGATAGA 455
QY 2098 aacttatgtgtataggaatgtgtacagcctttagccagagagtcatactatcaagac 2157
      |||||||
Db 456 AACTTATGTGTATAGGAGATGTGTACACCTTTTAGCCAGAGATCACTACTTATCAGAC 515
QY 2158 accaagagacacacacagctgtctgtcttctcaagcattgtcagataccaaagtggaga 2217
      |||||||
Db 516 ACCAGAGACACACACAGCTGTGCTGGCTTTTACAGCATTTGCTAGATACCAAGTGAGAG 575
QY 2218 catctgtgtgtatataatgataagactgtactgtgaagactgtatctccatccacctg 2277
      |||||||
Db 576 CATTCGTGTGTATGTATGTATGATGAGACTGTACTGTAGACTGTATCTCATCCACCTG 635
QY 2278 aagagaatctgtgtctcatcttcagagaagccctggc 2313
      |||||||
Db 636 AAGGAGAAATTGCTGCTCATTTTCAGAGAGCCCTGACC 671

RESULT 4
LOCUS AM245709 590 bp mRNA linear EST_07-JAN-2000
DEFINITION 2823003.5prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2823003 5',
      mRNA sequence.
ACCESSION AM245709
VERSION AM245709.1 GI:6588702
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
      National Institutes of Health, Mammalian Gene Collection (MGC)
      JOURNAL Unpublished (1999)
      COMMENT Other_ESTS: 2823003.3prime
      Contact: Robert Strausberg, Ph.D.
      Email: cgaabs-r@mail.nih.gov
      Tissue Procurement: DCTD/DBP cDNA Library Preparation: Ling
      Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
      Consortium (LNLN) DNA Sequencing by: Berkeley MGC sequencing
      Project Clone distribution: MGC clone distribution information can
      be found through the I.M.A.G.E. Consortium/LNLN at:
      www.bio.lnl.nih.gov/dbp/image/image.html Base Calling / Quality
      Trimming: PHRED from University of Washington Genome Center. Vector
      PHRAP suite. Poly-T identification: patchcut.pl from Berkeley
      Drosophila Genome Project. University of Washington Genome Center:
      http://www.genome.washington.edu
      Plate: LILCM10 row: 0 column: 4
      High quality sequence stop: 582.
      Location/Qualifiers
        1..590
        /organism="Homo sapiens"
```

```
/db_xref="taxon:9606"
/clone="IMAGE:2823003"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/ab_host="DH10B (phage-resistant)"
/Note="Organ: Lung; Vector: pORF7; Site:1: XhoI; Site:2:
ECORI; cDNA made by oligo-dT priming. Directionally
cloned into ECORI/XhoI sites using the following 5'
adaptor: GGCACAG(G). Size selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the Laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 192 a 127 c 161 g 110 t

ORIGIN

Query Match 21.0%; Score 581; DB 9; Length 590;
Best local similarity 100.0%; Pred. No. 8.7e-295;
Matches 581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 544 gggcaggaatccctccaccctgggaaacacactatccagaagatcagaacaacagcagg 603
      |||||||
Db 10 GGGCAGGAATCTCTCCACCTGGGAAAACATTCACAGATCCAGAAACAAACACACAGG 69
QY 604 atccattctgtcttagtgcaaaagcagaatgtatcagaaggagagaagctccagactcc 663
      |||||||
Db 70 ATCCATTCTGCTTTAGTGGCAAAAGCAGATGTGATTCAGAGGAGAGACCTCCAGACTCC 129
QY 664 tgtttggagagtaagcaaaatgtgcacttcaaaagcacttccagccaccctgaaagac 723
      |||||||
Db 130 TGTTTGGAGAGTAAACCAAAATGGCATTCAAGGCACTTTCAGACCCACCTGAAAGAC 189
QY 724 aacagccagacacagtcacaaagagaacacacagtggtgtatataagtggtccagctgaac 783
      |||||||
Db 190 AACAGCCACACACAGTCCAGGAGAAACACACAGTGTGATATAGGCTCCACCCCTGAAC 249
QY 784 ggaaggcagatcctagaaggaacagaacaaagtattgcatggttaagaagctcagaatttg 843
      |||||||
Db 250 GGAAGGCCATCTAGAGGAAACAGACAAAGTATGTGATGTAGAGTCTCAGGATTTGG 309
QY 844 gagaatcaaatatagaagagtggtggccaggtcttaacaggaagtcacaaacctcttagcc 903
      |||||||
Db 310 GAGAAATCAAAATATGAAGAGTTTGGGCCAGGCTTTATCAAGAGGTCAAACTCCTTAGCC 369
QY 904 tccagaagaacaaactggggagacaccttacaatgacatgagtggtggagacagcttg 963
      |||||||
Db 370 TCCAGAGACACAAACCTGGGAGACACTTACATGTACACTGAGTGGGAGACAGCTTTTG 429
QY 964 gcagtatgcagtcctcctcaaaaacccaagagacacactctggggaaagccttatgtgt 1023
      |||||||
Db 430 GCAGTATGTCACTCCTCATCAAAAACCCAAAGACACACTCTGGGGGAAAGCCTTATGTGT 489
QY 1024 gcaaggagtgctggcgaaggctttagcttggaatcacaacctgtatcacatcagaagacac 1083
      |||||||
Db 490 GCAGGGAATGTGGGCGAGGCTTTAGCTGGAAGTCAAACTGTATCAGATCAGAGAGACAC 549
QY 1084 actcaggggagaacacttatgtgtgcaagatgttggaga 1124
      |||||||
Db 550 ACTCAGGGGAGAAACCTTATGTGTGCAAGGATTTGTGACGA 590

RESULT 5
LOCUS BM449478 1009 bp mRNA linear EST_05-FEB-2002
DEFINITION AGENCOURT_6401009 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5493685
      5', mRNA sequence.
ACCESSION BM449478
VERSION BM449478.1 GI:18498518
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
```

REFERENCE 1 (bases 1 to 1009)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: L1AM12117 row: m column: 14
 High quality sequence stop: 649.

FEATURES
 source
 Location/Qualifiers
 1..1009

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5493685"
 /clone_lib="NHL_MGC_67"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."
 289 a 235 c 259 g 224 t 2 others

BASE COUNT
 ORIGIN
 Query Match 19.7%; Score 545; DB 10; Length 1009;
 Best Local Similarity 99.8%; Pred. No. 9.4e-276;
 Matches 595; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 206 ggcgtctgtgcatccagagatgtgctgtgtacttcaccagagagatgtgtgt 265
 |||||
 Db 196 ggcgtctgtgcatccagagatgtgctgtgtacttcaccagagagatgtgtgt 255
 OY 266 gaggccctctcagagacccttcacagaggagtgtgtctgtgagactataaccatctgt 325
 |||||
 Db 256 GAGCCCTGTCAGAGACCCTGCACAGGAGGTGATGTGAGACTTATTAACCATCTGCT 315
 OY 326 ctccactggaattccatctctaaacccaactcattgtctgcctgtgagcgagggaac 385
 |||||
 Db 316 CTCACCTGGAATTCATCTTAAACCAAACTCATGCTGCTGACGTGAGCGAGGGGAAC 375
 OY 386 gccctggaagagagagaaatgtccacttgaccctctgcagaaatcgaaagcagaat 445
 |||||
 Db 376 GCCCTGAGAGAGAGAGAAATGTCCACTGACCTCTCTCCAGATCAACACCAAAAT 435
 OY 446 tcaacttagtcctcctcctcctctgatttctccagtcagcaagctctcaagccaatgt 505
 |||||
 Db 436 TCAACTTAGTCCCTCCTGCTGCTGATTTTCTCAGTCAAGCAAGCTCTCAGCAACATGT 495
 OY 506 gtggtctgtatctctctcactgttttcaagttaatggaagaaatccctctcact 565
 |||||
 Db 496 GTGGCTGAGTCACTCTCTCAGCTCTTTTCAAGTTATGGGAGAAATCTCTCCACT 555
 OY 566 gggaaacactatccagaagatcagaacaacagcagatcattctgtcttagtgcga 625
 |||||
 Db 556 GGGAAACACTTCCAGAAATCAGAAACACAGCAGATTCATTCGTTAGTTCGCA 615
 OY 626 agcagaatgattcaagagagagaaactcagactcctgtttggagagtaagcaaaaa 685
 |||||
 Db 616 ACCAAGATGATTCAGAGGAGAGACTCCAGACTCTGTGGGAGAGTAAGCAAAAA 675
 OY 686 tggacttcaagagacttcagaccacttaaaagaaacagccagcagcatccaaga 745
 |||||
 Db 676 TGGCACTTAAAGGCACTTTCAGCCCACTGAGAACACAGCCGACAGCTCAAGGA 735

OY 746 agacacacagtgatgatagatggtccagccctgaacgagagagatctagagg 801
 |||||
 Db 726 AGACNACAGTGTGTGATATAGGTCCAGCCCTGACAGGAGGCAATCTAGAG 791

RESULT 6
 A1669533 563 bp mRNA linear EST 14-MAY-1999
 LOCUS wb88h09.x1 NCI_CGAP_Pri28 Homo sapiens cDNA clone IMAGE:2312801 3',
 DEFINITION mRNA sequence.
 ACCESSION A1669533
 VERSION A1669533.1 GI:4834307
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 563)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 www.bio.lnl.gov/bbrp/image/image.html
 Seg primer: -400p from Gibco
 High quality sequence stop: 454.
 Location/Qualifiers
 1..563

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="2312801"
 /clone_lib="NCI_CGAP_Pri28"
 /sex="male"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Organ: prostate; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI_CGAP_Pri28 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 985608-986759, 1101192-1101959, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaldo."
 BASE COUNT 162 a 125 c 150 g 126 t

ORIGIN
 Query Match 18.4%; Score 510; DB 9; Length 563;
 Best Local Similarity 99.8%; Pred. No. 2.5e-257;
 Matches 560; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2212 tggagacattctgtgtgatataatgcatgagactgtactgtgaagactgtatccatc 2271
 |||||
 Db 563 TGAGACATTGTGTGTGATATGATGATGAGACTGTACTGTATGATGATCTCAGC 504
 OY 2272 cactgaaagagaattgtcgtgtcatatttcaggaagccctgccttccactgtgtatg 2331
 |||||
 Db 503 CACCTGAAGAGAAATTCGCGCTCATATTCAGAGCCCTGCTCTCAGCTGTGATGG 444
 OY 2332 tgggtgtgaaacccggtcaggttaatgatagtgagagagcagtcgaatgccagaga 2391
 |||||
 Db 443 TGGGTGTGGAACCCGGTCAAGTATGATAGTGCAGAGAGCAGTCAATGCCAGGCA 384
 OY 2392 gataaggtgtgttacctgtgaaccacctaagctgaagagacagtcgcggttaatc 2451

```
|||||
Db 383 GATAGGGGTGCTGACCGTGGAACCCAAACCTTAAAGCTGAAGACAGTCCGGCTAATC 324
QY 2452 ctaatactgaatgagaacactgctctccatttggtgtgttcttcctgattatcccaa 2511
Db 323 CTATATGATGATGAGAACCTGTCTTCCATTTGGTGTCTTCCGATTAATCCCA 264
QY 2512 cccctcaactatttcaagtaacactgaccttccctaattggttttcaactgtgtccc 2571
Db 263 CCTTACACCTATTATTAGTATACCTGGCCCTTTCCTAATTGGTTTATACATGCTGCC 204
QY 2572 accctttgagtggtgcttcttgcaacttacaactcaagtcagtgtaattccctattctg 2631
Db 203 ACCTTTGAATGTGCTGCTTGCATCTACTTACAAATCAAGTCAACGTGATTCCTTATTCG 144
QY 2632 agcccaataaagaccagactgactgcaagtgaagagaagaatcaacctgtctgtgaagt 2691
Db 143 AGCCCATAAAGACCCAGACTGACGTGAGAGAGAAATCACCTGTGTGGGGGT 84
QY 2692 tggggaaccactccctgcatccctctccactgagagctgttcttctgctcaataaattc 2751
Db 83 TGGGGAACACTCCCTGATCCCTGCTGACTGAGAGCTGTCTTTCCTCAATAAATTC 24
QY 2752 tttcttaaccatccctcaacct 2772
Db 23 TTTTCTAACCATCTCACCCT 3

RESULT 7
AL047439 535 bp mRNA linear EST 29-FEB-2000
LOCUS DKFP586I0320.r1 586 (synonym: hute1) Homo sapiens cDNA clone
DEFINITION DKFP586I0320, mRNA sequence.
ACCESSION AL047439
VERSION AL047439.1 GI:4727354
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 535)
AUTHORS Mambutt,R., Heubner,D., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE EST (Mambutt, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Mambutt R
MIPS
Am Klopferstritz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sl sequence available.
This clone (DKFP586I0320) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, Germany; Email: clone@rzd.de.
FEATURES
source
1..535
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFP586I0320"
/clone_1lb="586 (synonym: hute1)"
/tissue_type="uterus"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI/MluI"
BASE COUNT 159 a 128 c 143 g 105 t
ORIGIN
Query Match 18.4%; Score 509; DB 9; Length 535;
Best Local Similarity 100.0%; Fred. No. 8.2e-257;
Matches 509; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1071 catcagaagacacactcaaggagaaaccttatgtgtgcaaggattgtgacaggctt 1130
Db 1 CATCAGAAGACACACATCAGGGGAGAAACCTTATGTGTGCAGAGATTGTGACAGAGCTTT 60
QY 1131 acttgaagtcgaacctcttaccatcagcagacacactcaaggctcaagcttatgtg 1190
Db 61 ACTTGAAGATCGAAGCTTTTACATCAGCGGACACACTCAGAGGCTCAAGCCTTATGTG 120
QY 1191 tgaagaagtatgtggaaggctttagcctgaagtcgaaccttatcccaaggaggcg 1250
Db 121 TGCAGGAATGTGGGAGAGCTTTACCTGCAACTCAACCTCAATACCCAGAGGGCG 180
QY 1251 cacacgaggagaaagccttatgttctcagggaatgtgagcgctgtcttcgcagcatca 1310
Db 181 CACACTGGGAGAGAGCTTATGTTCAGGAGATGTGGCGTGTGCTTCCACGACTTCA 240
QY 1311 cactgtgtcagacacaagagacacatcgaagagaagccttatctgcaaggagtg 1370
Db 241 CACCTGTCTAGACACAAGAGACACATTCAGAGAGAACCTTACATTTGCAGGAGTGT 300
QY 1371 gacagaagccttagcagaagtcacacactcatcagacacttaaggacacacagaagag 1430
Db 301 GAGCAAGGCTTTAGCCAGAGTCACTCATCAGACTTAAGACACACACAGAGAGAG 360
QY 1431 aagccttatgtatgcacagaatgtggtcgtcactttagctgtgaaatcaaacctcaaa 1490
Db 361 AAGCCTTATGTATGCACAGAGATGTGGCGTCATTAGCTGGAATCAAAACCTCAAAACA 420
QY 1491 caccagaagacacacatcgaagggttaaaccttatgttgcctgtgagtgcgagtgctt 1550
Db 421 CACCAAGGACACACATCAGGGGTAAACCTTATGTCTGCTGAGTGGCGGAGTCTTT 480
QY 1551 agcctgaagtcgaaccttaacaacacca 1579
Db 481 AGCCTGATGATCAACCTTAACAACACCA 509

RESULT 8
AI802142 558 bp mRNA linear EST 16-DEC-1999
LOCUS tx30c05.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2271080 3',
DEFINITION mRNA sequence.
ACCESSION AI802142
VERSION AI802142.1 GI:5367614
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 558)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA library Preparation: M. Bento Soares, Ph.D.
cDNA library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ULNL at:
www.bio.lnl.gov/bhrp/image/image.html
Insert Length: 771 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 455.
FEATURES
source
1..558
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2271080"
```


RESULT 10
BE178716 624 bp mRNA linear EST 22-JUN-2000
LOCUS PM4-HT0606-030400-001-cl0 HT0606 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE178716
ACCESSION BE178716.1 GI:8657868
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 624)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.V.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=est&PM4-HT0606-030400-001-cl0&ts=2000-04-03&td=1)
400-001-cl0&ts=2000-04-03&td=1)
Seq primer: puc 18 forward
High quality sequence start: 9
High quality sequence stop: 603.
Location/Qualifiers
FEATURES
source
1. 624
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0606"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESNES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 188 a 137 c 177 g 122 t
ORIGIN

Query Match 17.0%; Score 471; DB 9; Length 624;
Best Local Similarity 99.7%; Pred. No. 9.4e-237;
Matches 571; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 734 acagtcacgaagaacacacagtcgtgatagataggtccagccctgaacgagggcaga 733
|||||
DB 9 ACAGTCCACGAGACACACAGTGTGATATAGGCTCCAGCCCTGAAGGAGGGCAGA 68
|||||
QY 794 tcagaaggaacagacaaatattcagtggttagaagtcagagattggagagaatcaa 853
|||||
DB 69 TCTAGAGGAACACGACAAATATTGCTATGCTTGAAGTCTCAGGATTTGGAGAAATCAA 128
|||||
QY 854 atataagagatttggccagcttatacaagagtcacaaacctcttagctcagaagac 913

|||||
DB 129 ATATGAGACTTGGGCGACGCTTATCAAGAGTCAAACTCTTAGCTCCAGGAAGAC 188
QY 914 acaaacgtggggaacacacttaactgtacactggttgaggagacagcttggcagtatgc 973
|||||
DB 189 ACAAACTGGGAGACACCTTACATGATCACTGAGTGGGAGACAGCTTGGCAGTATGTC 248
QY 974 agtcctatcaaaaacccaagagacacactctgggggaaacccctatgtgtcagggaatg 1033
DB 249 AGTCTTATCAAAAACCAAGACACACTCTGGGGGAAACCTTATGTGTGCAAGGAGATG 308
QY 1034 tggcgagagcttaccgttggaagtcacacactgatacacatcagaagagacacacacag 1093
|||||
DB 309 TGGGCGAGGCTTTACGCGAAGTCAAACCTGATCACACACACAGACACACTCAGGGGA 368
QY 1094 gaacacttgatgtggaagagattgtggaagagcttacttggaagtcagacactttac 1153
DB 369 GAAACCTTATGTGTGCAAGGATTTGGACGAGGCTTACTTGAAGTCGAACCTCCTTAC 428
QY 1154 acatcagcgagacacactcagagctcactatgtgtgcaagaaatgtgagagagctt 1213
DB 429 ACATCAGCGGACACACACTCAGGCTCAAGCCTTATGTGTGCAAGGAATGTGGCAGACTT 488
QY 1214 tagcctgaagtcacacacttaaccacagagggcgacacttggggagagccttatgt 1273
DB 489 TAGCCTGAAGTCAAACTCATTACCCACAGAGGGCGCACACTGGGAGAAAGCCTTATGT 548
QY 1274 ttgcaggaagtgtggtggtgtggttgcagca 1306
DB 549 TTGCAGGGAATGTGGCGGTGCTTCCGACGCA 581

RESULT 11
BE111754/c 512 bp mRNA linear EST 20-OCT-2000
LOCUS BE111754
DEFINITION IMAG:3523123 3', mRNA sequence.
ACCESSION BE111754
VERSION BE111754.1 GI:10941444
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 512)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps@mail.nih.gov
This clone is available royalty-free through LNL; contact the IMAG Consortium (info@lmage.llnl.gov) for further information.
Seq primer: -40UP from Glibco
High quality sequence stop: 447.
Location/Qualifiers
FEATURES
source
1. 512
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAG:3523123"
/clone_lib="Soares_NSF_F8_9W_OT_PA_P_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pTZ19D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; Equal amounts of plasmid DNA from five normalized libraries were mixed, and ss circles were made in vitro. Following NHP purification, this DNA was used as tracer in a subtraactive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and clones: Soares NB2HP pool 1: 309384-310919, 323208-325895 Soares NB2HP pool 1: 145032-147335, 147720-148103, 148872-149255, 15002 -

150407, 151176-152327 Soares ND2HF8-9W pool 1:
 758280-760583, 772104-774407 Soares NBHPA pool 1:
 304776-306311, 320136-322823, 326280-326663 Soares NBHOF
 pool 1: 723720-726407, 739080-740999 Subtraction by Bento
 Soares and M. Fatima Bonaldo."

BASE COUNT 144 a 111 c 140 g 117 t
 ORIGIN

Query Match 16.5%; Score 458; DB 10; Length 512;
 Best Local Similarity 99.8%; Pred. No. 6.6e-230;
 Matches 508; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2260 tctatctccatccaccctgaaggaaattgctggtccatttttcaggagcccttcctc 2319
 |||||||
 DB 512 tctatctccatccaccctgaaggaaattgctggtccatttttcaggagcccttcctc 453
 |||||||

OY 2330 cactgtgagatggttggttgatgaaccgcgtcagtaataatgagtgagagagcagta 2379
 |||||||
 DB 452 cactgtgagatggttggttgatgaaccgcgtcagtaataatgagtgagagagcagta 393
 |||||||

OY 2380 aatgcccagagcagataggggtgtgtaccgtgtgaaacccaacttaagctgaagacgt 2439
 |||||||
 DB 392 aatgcccagagcagataggggtgtgtaccgtgtgaaacccaacttaagctgaagacgt 333
 |||||||

OY 2440 cccggcctaactccatactgaattgagaacctgtctccatttggtgtgtcttcctcc 2499
 |||||||
 DB 332 cccggcctaactccatactgaattgagaacctgtctccatttggtgtgtcttcctcc 273
 |||||||

OY 2500 gatgtatcccaacccttcacctaatttaacgtacacctgaccttccttaattggttttac 2559
 |||||||
 DB 272 gatgtatcccaacccttcacctaatttaacgtacacctgaccttccttaattggttttac 213
 |||||||

OY 2560 actgtctgtgccaccttttgatgtgtgctctttgcatacttaacaatcagctcaacgtgtat 2619
 |||||||
 DB 212 actgtctgtgccaccttttgatgtgtgctctttgcatacttaacaatcagctcaacgtgtat 153
 |||||||

OY 2620 tcccatctctgaagcccaataaaagaccagactcagctgagtgagagagaataacacc 2679
 |||||||
 DB 152 tcccatctctgaagcccaataaaagaccagactcagctgagtgagagagaataacacc 93
 |||||||

OY 2680 tgcctgtgagaggttgaggaccactccctgcatccctccctcaactgagagctgtctcttcg 2739
 |||||||
 DB 92 tgcctgtgagaggttgaggaccactccctgcatccctccctcaactgagagctgtctcttcg 33
 |||||||

OY 2740 tcaataaaattctttctaccatcctca 2768
 |||||||
 DB 32 tcaataaaattctttctaccatcctca 4

RESULT 12
 BE463501 505 bp mRNA linear EST 27-JUL-2000
 LOCUS hw24d02.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3183843 3',
 DEFINITION mRNA sequence.
 ACCESSION BE463501
 VERSION BE463501.1 GI:9509274
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 505)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILN, send email to:
 info@image.llnl.gov
 Seq primer: -400P from Gibco
 High quality sequence stop: 455.
 Location/Qualifiers

FEATURES
 source
 1..505
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3183843"
 /clone_lib="NCI_CGAP_Kid11"
 /lab_host="DH10B"
 /note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
 a modified polylinker. Site 1: Not 1; Site 2: Eco RI;
 Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
 prepared, and ss circles were made in vitro. Following HAP
 purification, this DNA was used as tracer in a subtractive
 hybridization reaction. The driver was PCR-amplified cDNAs
 from a pool of 5,000 clones made from the same library
 (cloneids 1322376-1323911, 1456007-1456775, and
 1500552-1502855). Subtraction by Bento Soares and M.
 Fatima Bonaldo."

BASE COUNT 142 a 110 c 140 g 113 t
 ORIGIN

Query Match 16.4%; Score 454; DB 10; Length 505;
 Best Local Similarity 99.8%; Pred. No. 8.4e-228;
 Matches 504; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2263 atctcatccacctgaaggaaattgctgctatcttcaggagcccttcctccac 2322
 |||||||
 DB 505 atctcatccacctgaaggaaattgctgctatcttcaggagcccttcctccac 446
 |||||||

OY 2323 tgtgtatggttggttgatgaaccgcgtcagtaataatgagtgagagagcagtaaat 2382
 |||||||
 DB 445 tgtgtatggttggttgatgaaccgcgtcagtaataatgagtgagagagcagtaaat 386
 |||||||

OY 2383 gcccaagcagataggggtgtgtaccgtgtgaaacccaacttaagctgaagcagctccc 2442
 |||||||
 DB 385 gcccaagcagataggggtgtgtaccgtgtgaaacccaacttaagctgaagcagctccc 326
 |||||||

OY 2443 ggcataacctcaactgaattagaaacctgtctccatttggtgtcttcctccgat 2502
 |||||||
 DB 325 ggcataacctcaactgaattagaaacctgtctccatttggtgtcttcctccgat 266
 |||||||

OY 2503 tgaaccacaccttcacctaatttaacgtacacctgaccttccttaattggttttaacct 2562
 |||||||
 DB 265 tgaaccacaccttcacctaatttaacgtacacctgaccttccttaattggttttaacct 206
 |||||||

OY 2563 gctgtgccacaccttttgatgtgtgctcttgcatcttaagaatcagctcaacgtgtatttc 2622
 |||||||
 DB 205 gctgtgccacaccttttgatgtgtgctcttgcatcttaagaatcagctcaacgtgtatttc 146
 |||||||

OY 2623 cctattctgagcccaataaaagaccagactcagctgagtgagagagaataacacctgc 2682
 |||||||
 DB 145 cctattctgagcccaataaaagaccagactcagctgagtgagagagaataacacctgc 86
 |||||||

OY 2683 tgtgagaggttgaggaccactccctgcatccctccctcaactgagagctgtctcttcgcta 2742
 |||||||
 DB 85 tgtgagaggttgaggaccactccctgcatccctccctcaactgagagctgtctcttcgcta 26
 |||||||

OY 2743 ataaattctttctaccatcctc 2767
 |||||||
 DB 25 ataaattctttctaccatcctc 1

RESULT 13
 AM401692 486 bp mRNA linear EST 16-FEB-2000
 LOCUS AM401692
 DEFINITION UI-HF-BKO-aaf-c-03-0-UI.r1 NIH_MGC_36 Homo sapiens cDNA clone
 IMAGE:3053549 5', mRNA sequence.

ACCESSION AM401692
VERSION AM401692.1 GI:6920378
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 486)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps@remail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: M.B. Soares Lab
CDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/BLNT at:
www.bio.livl.gov/brp/image/image.html
Seq primer: M13 Forward
Location/Qualifiers
1..486
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3053549"
/clone_lib="NIH_MGC_36"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/cell_line="MGC85"
/lab_host="DH10B (LT1)"
/note="Vector: pT7m3-Pac; Site.1: NotI; Site.2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(0.5-1.5kb). Directionally cloned. Cells provided by Louis
M. Staudt, Ph.D. Library preparation by Maria de Fatima
Bonaldio, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT 138 a 113 c 136 g 98 t 1 others
ORIGIN

Query Match 16.2%; Score 449; DB 9; Length 486;
Best Local Similarity 100.0%; Pred. No. 3.6e-225;
Matches 449; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 963 ggcagatgtcagtcctccatcaaaaacccaagacacactctggggaaagccttaagt 1022
|||||
Db 8 GGCAGATGTGTCAGTCTCCATCAAAAACCCCAAGACACACTCTGGGGAAAGCCTTATGTG 67

QY 1023 tgcagggaatgtggcgaagccttaagctgaagtcacaacctgaacaacacacacagaagaca 1082
|||||
Db 68 TGCAGGGAATGTGGGAGGCTTACGTGAGATCAAACTGATACACATCAGAGAGACA 127

QY 1083 caatcaggggagaacacctatgtgtcgaagatgtggaagagccttaacttggaagtcg 1142
|||||
Db 128 CACTCAGGGGAGAAACCTTATGTGTGCAAGAGATTGTGAGAGAGCTTTACTTGAATCG 187

QY 1143 aacctcttaacatcagcgacacactcagggtcgaagccttatgtgtcgaagaaatgt 1202
|||||
Db 188 AACCTCTTTACATCAGCGGACACACTCAGGGCTCAAGCCTTATGTGTCAAGGAATGT 247

QY 1203 gggcagaagccttaagcctgaagtcacaacctcaatcaacccaagaggcgacacactggag 1262
|||||
Db 248 GGGCAGAGCTTTAGCCGTAAGAGTCAAACTCATTAACCCACAGAGGGGCGACACTGGGGAG 307

QY 1263 aagccttaagcttcgaaggaaatgtgggtgtgtccttcgcagacatcaacactgtgtcaga 1322
|||||
Db 308 AACCCCTTATGTGTGCAAGGGAATGTGGCGGTGGCTTTCGCAGCATTTACACACTGTGTGAGA 367

QY 1323 caaagaagacacatcagaagagaagccttaacttgacagggaagtgtgaagccttt 1382
|||||
Db 368 CACAAGAGCACATTCAGAGAGAAAGCCTTACATTTGACGAGGAGTGTGAGCAAGGCTTT 427

QY 1383 agccagaagtcacacccatcacagacact 1411
|||||
Db 428 AGCCAGAAAGTCACACCTCATCAGCACTT 456

RESULT 14
LOCUS AV720986 486 bp mRNA linear EST 16-OCT-2000
DEFINITION AV720986 HTB Homo sapiens cDNA clone HTB1C05 5', mRNA sequence.
ACCESSION AV720986
VERSION AV720986.1 GI:10818138
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 486)
AUTHORS Gu, Y., Peng, Y., Song, H., Huang, Q., Yang, Y., Gao, G., Xiao, H., Xu, X.,
Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu,
S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R.,
Chen, J., Chen, Z., and Han, Z.
Homo sapiens CDNA HTB clones
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
1..486
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HTB1C05"
/clone_lib="HTB"
/tissue_type="Hypothalamus"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site.1: EcoRI; Site.2:
XhoI"

BASE COUNT 104 a 134 c 110 g 137 t 1 others
ORIGIN

Query Match 15.1%; Score 419; DB 9; Length 486;
Best Local Similarity 100.0%; Pred. No. 2.4e-209;
Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2269 atccacctgaagagaatgtctgtcattcttaaggaagcccttcctcactgtgga 2328
|||||
Db 1 ATCCACTGAAAGAGAAATTCCTGCTCATTTTATGAGAGCCCTCCTCTCATCTGTGGA 60

QY 2329 tggctgggtctgtgaaacccggtcaggttaatgatatgtgcagagagagtcacaaatgccag 2388
|||||
Db 61 TGGTGGGTGTGGAAACCCGCTCAGGTAATGATGATGTGCGAGGAGCAATGATGCCAG 120

QY 2389 gtcagatagggtgtgtgtaacctgtgtgaaacccaaccttaagctgaagaacagtcaggctaa 2448
|||||
Db 121 GCAGATPAGGGGTGGTACCTGGTGAACCCAACTTAAAGCTTAAGACAGCACTCCGGCTAA 180

QY 2449 atcctcatctgaatgagaacctgtctccatcttgatgtgtgtcttcctccagatgac 2508
|||||
Db 181 ATCCTCATCTGAAATGTGAGAACTGTCTCCATTGGTGCTGTTCCTCCGATTGATG 240

QY 2509 caacctcaactatttaccgtataacctgtcccttccttaattggttttaacactgtgtg 2568
|||||
Db 241 CAACCCCTTACCTAATTTTACGTATACCTGCCCTTCTAATGTTGTTTACACTGCTG 300

QY 2569 ccacactttgagtggtgtcctttgcatacttaacaatcagtcgaagtgatctccctatt 2628
|||||
Db 301 CCCACCTTTTGAAGTGTGCTTTGTGATACATTAAATCAAGTGATTTCCCTTATT 360

Qy 2629 ctgagccataaagaccagactcagctgcaatgaggaagaataccctgtctgg 2687
|||||
|||
Db 361 ctgagccataaagaccagactcagctgcaatgaggaagaagaattaccctgtctgg 419

[illegible]

TITLE	Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
JOURNAL	Unpublished (2000)
COMMENT	Contact: John Quackenbush The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 3528 Fax: 301 838 0208 Email: johngettigr.org Plate: 207 Seq primer: Forward.

[illegible][illegible]

RESULT	16
LOCUS	BE391140
DEFINITION	697 bp mRNA linear EST 21-JUL-2000
ACCESSION	BE391140
VERSION	601286705.F1 NIH_MGC_44 Homo sapiens CDNA clone IMAGE:3613513 5',
KEYWORDS	mRNA sequence.
SOURCE	BE391140
	BE391140.1 GI:9336505
	Est.
	human.

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Plate: L1CM274 row: 1 column: 02
High quality sequence stop: 635.

FEATURES	source	location/Qualifiers
	1..697	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:3613513"	
	/clone_id="NH.MGC.44"	
	/tissue.type="endometrium, adenocarcinoma cell line"	
	/lab_host="DH10B (phage-resistant)"	
	/note="Organ: uterus; Vector: pOT7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."	
BASE COUNT	185 a	171 c 201 g 140 t
ORIGIN		
Query Match	14.9%;	Score 413; DB 10; Length 697;
Best Local Similarity	99.6%;	Pred. No. 3.8e-206;
Matches 513; Conservative	0; Mismatches 2; Indels	0; Gaps 0;
QY 1280	ggaatgtggggttgcttcgcgcagcatctacacctgtgcagacaagaagacacatc	1339
DB 1	GGAAATGTGGGGCTTCGTCACACATTCACACCTGTGTCAACACAAGAGACACATGC	60
QY 1340	aagagagaagccttacattgcaggagatgtgaagcaagccttaagccgaagtacacct	1399
DB 61	AGGAGGAAGCCTTACATTTGCAGGAGAGTGTAGACAAAGCCTTTAGCCACAACTCACACTT	120
QY 1400	catacgaaccttaagaaacacacacaggaagaagccttaagtatcaagaagtgtgggct	1459
DB 121	CATCAGACACTTAAGGACACACACAGAGAGAAAGCCTTATGTATGCACAGAAATGTGGCG	180
QY 1460	tcaacttaagtgaataccaacctcaaaacacacacagagagacacactcagggttaaac	1519
DB 181	TCACCTTACTCTGGAATCAAACTCAAAACACACACGAGACACACACTCGGGGTTAAACC	240
QY 1520	ttaatgtcgtcgtgagtcgcggcgagtgcttaagcctgaaglcataaccttaacaaaccca	1579

```
|||||
Db 241 TTTATGCTGCTGAGTGGCGGAGTGTCTTACCTGAACTCAAACTTAAACAACACCA 300
QY 1580 gaagtcacacacgggggaaagaccatttatagtacggagtggtggcgaagcttaccgg 1639
Db 301 GAGGTACACACACGGGGGAGAGACATTGTATGTACGAGATGGGGGAGGCTTTACCCG 360
QY 1640 gaaatcaacctgatacagcaccagagacacactcagggaggaagccatttgatgtgc 1699
Db 361 GAAATCAACCCCTGAGCAGCAGACAGACACTCAGGGGAGAGACGATTGTATGTGC 420
QY 1700 tgaagtgtgacagagcttataatgaatcaccctcattcaccacagagacacattc 1759
Db 421 TGAGTGTGGACGAGGCTTTAATATGATGTCACCCCTCATTTACACACAGACACATTC 480
QY 1760 aggggaaagcccttatgtgtcagggagtggtgca 1794
Db 481 AGGGGAAAAGCCTTTATGTGTGACGAGAGTGTGGCA 515

RESULT 17
LOCUS BG928720 673 bp mRNA linear EST 06-NOV-2001
DEFINITION HNC72-1-B2.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA
ACCESSION BG928720
VERSION BG928720.1 GI:14323243
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 673)
Kumar,S., Connor,J.R., Dadds,R.A., Halsey,W., Van Horn,M., Mao,J.,
Sathe,G., Mul,P., Agarwal,P., Badger,A.M., Lee,J.C., Gowen,M. and
Lark,M.W.
Identification and initial characterization of 5000 expressed
sequenced tags (ESTs) each from adult human normal and
osteoarthritic cartilage cDNA libraries
Osteoarthritis Cartilage 9 (7), 641-653 (2001)
21482651
JOURNAL COMMENT
MEDLINE Contact: Sanjay Kumar
UM2109
GlaxoSmithKline
709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
Tel: 610-270-7245
Fax: 610-270-5598
Email: sanjay.kumar-1@sk.com
Seq primer: 17.
FEATURES
Source
1..673
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HNC (Human Normal Cartilage)"
/tissue_type="cartilage"
/lab_host="E.coli DH10 B"
/note="Vector: pSPORT 1; Site_1: SalI; Site_2: NotI;
Directional"
BASE COUNT 181 a 156 c 189 g 145 t 2 others
ORIGIN
Query Match 14.6%; Score 405; DB 10; Length 673;
Best Local Similarity 99.5%; Pred. No. 6.3e-202;
Matches 555; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1775 tatgtcaaggagtggtgcagaaggttcgagcaagcctaactgtttaaggacaaag 1834
Db 236 TATGTCAAGGAGATGTGGCAGAAGTTTCGGCAGAAAGCTTAACCTTTATAGCACAAAG 295
QY 1835 ggcacacccaagtgtccttctgtgtcaggagtggtggcgaagccttctgtcagtaaac 1894
Db 296 GGCACACTAGAGTCCCTTTGTGTGTCAGGAGTGTGGCAAGGCTTTGTGCTTAAGTTAAC 355
QY 1895 tctcattaaccaccaagagacacagggggggaagcctaatgtgtcagggaagtgtg 1954
Db 356 TCTCATTAACACACAGAGACACACGAGGGGGAAGCCTCATGTGTGACGGAGGTGG 415
QY 1955 gcaagccttaagccggcgaclcacacctcattagacacagagagacatlcagagaaga 2014
Db 416 GCAAGCCTTAAAGCCGCGCAGCTCACACCTCAATTAGACACAGAGACACATTCAGAGAGAA 475
QY 2015 gccctatattgcagaagtgtagcgggctttagtcggaagtcacacctatcagaca 2074
Db 476 GCCTTATATTTTGCAGAAAGTGTGACGAGCAGGCTTTAGTCCGAAGTCCAACTTATCAGACA 535
QY 2075 tcaagagacacactcagagtagaacttatgtgtataaggaatgtgtacagccttag 2134
Db 536 TCAGAGGACACACTCAGATAGAACTNTATGTATAGGGAATGTGTACAGCCTTTAG 595
QY 2135 ccaggagtcatactcatalcacagacacagagacacacagatgtgtgttlttcaagc 2194
Db 596 CCAGAGTCATFACCTTCATCAGACACAGAGACACACAGTGTGTGCTTTTCAGGC 655
QY 2195 attgtcagatccaagt 2212
Db 656 ATTGCTAGATACCAAGT 673

RESULT 18
LOCUS AW513018 453 bp mRNA linear EST 03-MAR-2000
DEFINITION x176f03.x1 NCI-CGAP-Ut1 Homo sapiens cDNA clone IMAGE:2792381 3',
mRNA sequence.
ACCESSION AW513018
VERSION AW513018.1 GI:7151096
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 453)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL COMMENT
MEDLINE Contact: Robert Strausberg, Ph.D.
Email: cgapbs-ri@mail.nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/BLNI at:
Image.llnl.gov/image/html/lifresources.shtml
Seq primer: -400P from Gibco
High quality sequence stop: 423.
FEATURES
Source
1..453
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2792381"
/clone_lib="NCI-CGAP_Ut1"
/tissue_type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
```


JOURNAL Unpublished (1997)
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estwats@wustl.edu
 This clone is available royalty-free through LNLN; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Seq primer: -28m3 rev2 ET from Amersham
 High quality sequence stop: 428.

FEATURES
 source Location/Qualifiers
 1. 451
 /organism="Homo sapiens"
 /db_xref="GDB:598677"
 /db_xref="taxon:9606"
 /clone_image="IMAGE:789160"
 /clone_lib="Soares_total_fetus_Mb2HF8_9w"
 /dev_stage="8-9 weeks"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from pooled 8-9 week
 (total) fetus material with a Not I - oligo(dT) primer [5'
 TGTTACCAATCTGAAGTGGAGCGGCGCTTAATTTTATTTTATTTT 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 122 a 97 c 121 g 111 t
 ORIGIN

Query Match 14.2%; Score 395; DB 9; Length 451;
 Best Local Similarity 100.0%; Pred. No. 1.1e-196; Mismatches 0; Indels 0; Gaps 0;

Matches 395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2020 atatttcgaagaatgtcgacggcgtttagtcgaagtcacacccattatcacacacaga 2079
 Db 4 ATATTTCAGAAAGTGTGAGCGGGCTTATGTCGGAAGTCCAACTTATCAGACATCAGA 63
 QY 2080 ggaacactcaagatagaacttatgtgtataggaatgtgtacagcccttaagcaga 2139
 Db 64 GGACACACTCAGATGAAACTTATGTGTATAGGGAATGTGTACAGCCTTTAGCCAGG 123
 QY 2140 agtcatactcatcacacacacacacacacagtcgtgtgtctttttcaacattgc 2199
 Db 124 AGTCATCTTCATCAGACACACAGAGACACACAGAGCTGTGGCTTTTTCAGCCATTCG 183
 QY 2200 taataacaaagtgaacatctctgtgtgtatcatcatgaactgaactgaact 2259
 Db 184 TAAATACCAAAAGTGAAGACTTCTGTGTGATTATGCAATGACATGTAAGACT 243
 QY 2260 tttatctcatcacactgaagaatgtgtgtcatatttcagagccctgtccttcc 2319
 Db 244 TGTATCTTCATCCACCTGGAAGGAATGTGTGCTCATTTTTCAGAGCCCTGTCCCTCC 303
 QY 2320 caactgtgagtgtgtgtgtgtgtgaacccggtgcagtaagatgtgtgcaggaagcag 2379
 Db 304 CACTGTGAGTGTGTGTGTGTGTGAACCCGGTCCAGTAATGATAGTGGCAGAGGACAGTCA 363
 QY 2380 aatgcccaagcagatagaggtgtgtacctgtgtga 2414
 Db 364 AATGCCCAGGACAGATAGGGGTGTGTCTGTGTGA 398

RESULT 21 435 bp mRNA linear EST 27-AUG-1998
 AI081665/c
 LOCUS ou69e06.s1 NCI_CGAP_Br2 Homo sapiens cDNA clone IMAGE:1632514 3',
 DEFINITION mRNA sequence.
 ACCESSION AI081665

VERSION AI081665.1 GI:3418457
 EST.
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.

CDNA library preparation: M. Bento Soares, Ph.D.
 CDNA library arrayed by: Greg Lennon, Ph.D.
 DNA sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 www.bio.lnl.gov/db/rp/image/image.html
 Insert length: 862 Std Error: 0.00
 Seq primer: 40m3 fwd. ET from Amersham
 High quality sequence stop: 410.

FEATURES
 source Location/Qualifiers
 1. 435
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="IMAGE:1632514"
 /clone_lib="NCI_CGAP_Br2"
 /sex="female, pooled"
 /tissue_type="breast"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. 1st strand cDNA was prepared from pooled bulk
 breast tumor tissue, and was then primed with a Not I -
 oligo(dT) primer. Double-stranded cDNA was ligated to Eco
 RI adaptors (Pharmacia), digested with Not I and cloned
 into the Not I and Eco RI sites of the modified pT73
 vector. This library is the normalized version of
 NCI-CGAP Br1.1. Library was constructed by Bento Soares
 and M. Fatima Bonaldo."

BASE COUNT 122 a 90 c 122 g 101 t
 ORIGIN

Query Match 13.8%; Score 383; DB 9; Length 435;
 Best Local Similarity 99.8%; Pred. No. 2.3e-190; Mismatches 1; Indels 0; Gaps 0;
 Matches 433; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2339 tggaaacccggtcagtgatagtagtgcagagagcagtcacaaatgccagcagataggg 2398
 Db 435 TGGAAACCCGGTCAAGTATGATAGTGCAGAGGCAAGTCAATGCCAGGACAGATAGGG 376
 QY 2399 gtgggtacctgtgaaacccaacttaaaagctgaagaacagtcacgaataatccatc 2458
 Db 375 GTGGGTACTGTGGAAACCAACCTTAAGCTGAAAGACAGTCCCGCTAAATCCTCATAC 316
 QY 2459 tgaatgaaacctgtctcccatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2518
 Db 315 TGAATTTGAACACCTGTCTTCCATTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 256
 QY 2519 ccattttacgtatatacctcccttcctcatttggttttaactgtgtgtgtgtgtgtgt 2578
 Db 255 CCTATTTTACGTRACCTCTCCCTTCTTAATGTGTTTACACTGTGTGTGTGTGTGTGT 196
 QY 2579 ggtgtgtccttgcatacttaacaaatcagtcacagtgatcccatcctatcgtgagccat 2638
 Db 195 GAGTGTGCTTGTGATTAATCAATCAAGTCAAGTGTATTCCTTATTCAGAGCCAT 136
 QY 2639 aaaaagccagactcagctgcagtgagagagaatcacctcgtgtgtgtgtgtgtgtgt 2698
 Db 135 AAAAGACCCAGACTCAGTGCAGTGAAGAGAGAAATCACCTGTGTGTGTGTGTGTGTGT 76

Query Match 13.6%; Score 378; DB 10; Length 455;
Best Local Similarity 99.8%; Pred. No. 1e-187;
Matches 428; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2341 gaaccgcgtcaagatgatgatgagagagcagtcacaaatgccagagcagatagggt 2400
|||||
Db 429 GAACCCGGTCAGATGATGATGTCAGAGGACAGCAAAATCCAGAGAGATAGGGT 370
|||||
QY 2401 gggtagctggtaaacccaacttaagctgaagacagtcgcgcgaatccatcactg 2460
|||||
Db 369 GGGTACTGGTGAACCAACCTTAAGCTGAAGACAGTCCGGCTAAATCTCATACTG 310
|||||
QY 2461 aattgaacactgctctccatctggtgctgcttcctccgattgatacccaacctacc 2520
|||||
Db 309 AATTGAGAACCTGCTTCCCATTTGGTGCTTCCCTCCGATTGATGACCAACCTTCAC 250
|||||
QY 2521 tatttaagctataccgctccctccctccatctggttcttaacactgctgcccactttga 2580
|||||
Db 249 TATTTTACGTATACGTGCTTTCCTAATTGGTTTACACTGCTGCTGCTTTCGA 190
|||||
QY 2581 gtggtgcttgatactactcaaatcaatcaagctgctgcttctccctattctgagccataa 2640
|||||
Db 189 GTGGTGGCTTGGTACTTCAATCAATCAGTCAAGCTGATTCCTTATTCGAGCCATTA 130
|||||
QY 2641 aagaccacagctcagctcagtgagagagaaatcacacctgctgtgaggttggagagca 2700
|||||
Db 129 AAGACCCAGACTCAGCTGAGTGAAGAGAAATCAACCTGCTGTGGGGGTTGGGAGCA 70
|||||
QY 2701 ctccctgcacccctctccactgagagctgcttcttgcataataatctcttctacc 2760
|||||
Db 69 CTCCTGACATCCCTCTCCACAGAGCTGCTTCTTTCATTAATAATCTTTCAC 10
|||||
QY 2761 catccctac 2769
|||||
Db 9 CATCTCTAC 1

RESULT 24
BE746403 728 bp mRNA linear EST 15-SEP-2000
LOCUS BE746403
DEFINITION 601579525F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3528207 5',
mRNA sequence.
ACCESSION BE746403
VERSION BE746403.1 GI:10160395
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 728)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DPF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM759 row: a column: 08
High quality sequence stop: 726.
Location/Qualifiers
1..728
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3528207"
/clone.lib="NIH_MGC_9"
/tissue.type="adenocarcinoma cell line"
/lab.host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
BclRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 217 a 169 c 192 g 150 t
ORIGIN

Query Match 13.6%; Score 378; DB 10; Length 728;
Best Local Similarity 100.0%; Pred. No. 1e-187;
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1086 tcagggagaaacattatctgtaagagattgtgagcagagcttacttggaatgtaac 1145
|||||
Db 244 TCAGGGAGAAACCTTATGTGTCAAGATGTGACAGAGCTTACTTGGAGTGCAGAC 303
|||||
QY 1146 ctcttacaatcagcagcacactcaggtctcaagcctatgtgtgcaagaatgtgg 1205
|||||
Db 304 CTCTTACACATCAGCGGACACACTCAGGGCTCAACCTTATGTGTGCAAGATGTGG 363
|||||
QY 1206 cagagcttaagcctgaatcaaacctcaatcaccacagaaggcgacactggagaag 1265
|||||
Db 364 CAGAGCTTACGCTGAAGTCAAACTCATTTACCCACAGAGGCGCACACTGGGGAAG 423
|||||
QY 1266 cctatggttcagaggaatgtgaggtgtgcttgcgaacattcaacctgtcaaac 1325
|||||
Db 424 CCTTATGTTTGCAGGAAATGTGGGCTGTGCTTTCGACACTTCAACCTGTCACACAC 483
|||||
QY 1326 aagagacacatcagaagaagaagccttacttgcaaggagtgtgaagaagctttagc 1385
|||||
Db 484 AAGAGACACATTCAGAGAGAAAGCTTACATTTGCAAGGAGTGTGACAGAGCTTTAC 543
|||||
QY 1386 cagaagtcacacctcaatcagaacttaagagacacacagaagagagccttatgtac 1445
|||||
Db 544 CAGAACTCACACCTCATTCAGACACTTAAGACACACAGAGAGAAAGCTTATGTATGC 603
|||||
QY 1446 acagaatgtgagcgtcac 1463
|||||
Db 604 ACAGAAATGTGGGCTCAC 621

RESULT 25
A1016684 424 bp mRNA linear EST 27-AUG-1998
LOCUS A1016684/C
DEFINITION ou96d05.x1 NCI_CGAP_Kid3 Homo sapiens cDNA clone IMAGE:1635657 3',
mRNA sequence.
ACCESSION A1016684
VERSION A1016684.1 GI:3231020
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 424)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bto.llnl.gov/btrp/image/image.html
Insert Length: 1136 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 392.

FEATURES
source
Location/Qualifiers
1. .424
/organism="Homo sapiens"
/db.xref="taxon:9606"
/clone="IMAGE:1635657"
/clone_lib="NCI-CGAP_K1d3"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT7p3D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer, double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7p3 vector. mRNA source: 2 pooled kidneys. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 119 a 88 c 118 g 99 t
ORIGIN

Query Match 13.4%; Score 372; DB 9; Length 424;
Best Local Similarity 99.8%; Pred. No. 1.5e-184;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2346 ccggtcaggaatagatagtcaggaagcagtcacaatgccagcagatagagtgagta 2405
|||||
DB 424 CCGGTCAAGTATGATGATGCGAGGAGGAGCAATGCCAGAGATAGGGGTGGGTA 365
|||||

QY 2406 cctgtggaacccaaccttaaacgtgaagacagtcgcggcctaaatccctactgaatg 2465
|||||
DB 364 CCGGTGAACCCCAACCTTAAAGCTGAAGACAGTCCGGCTAAATCCATATCTAATG 305
|||||

QY 2466 agaaccctctcccatctgtgtcttcctcccgatgatcccaacctcaactatct 2525
|||||
DB 304 AGAACCTCTCTCCCATTTGGGTGCTTCTCCGATGATGCCAACCTTACCTATT 245
|||||

QY 2526 taagataacccctccctccctaaatgttttacactgtctgtccacactttgaatgt 2585
|||||
DB 244 TACGATATACCTGCGCTTCTCTAATGTTTATACACTGCTGCGCCACTTTTGAATG 185
|||||

QY 2586 gcccttgatacttacaatcagtcagtgatccctatctcgagccctataaaagac 2645
|||||
DB 184 GCCCTTGATCTTAAACAATCAGTCAACGTGATTCCTTATTCGACCCCTAAAGAC 125
|||||

QY 2646 ccaagcctagctcagtcagtcaggaagaatccaccctcgtgtgaggttgaggacaccc 2705
|||||
DB 124 CCAAGCTCAGCTGCGAGTGAAGAGAAATCAACCTGCTGTGGGGTGGGAGCACTCC 65
|||||

QY 2706 tgcataccctccacagtcagagctgtctcttgcataaataatctcttaccatcc 2765
|||||
DB 64 TGCATCCCTCTCCACTGAGAGCTGTCTTTGCTCAATAAATCTTTTACCATCC 5
|||||

QY 2766 tca 2768
|||||
DB 4 TCA 2

RESULT 26
LOCUS A1440109 428 bp mRNA linear EST 30-MAR-1999
DEFINITION U156h06.x1 NCI-CGAP_Lym12 Homo sapiens cDNA clone IMAGE:2134523 3',
mRNA sequence.
ACCESSION A1440109
VERSION A1440109.1 GI:4308857
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 428)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL
COMMENT
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapb-jemail.nih.gov
Life Technologies catalog #: 11547-015
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrrp/image/image.html
Insert length: 1031 Std Error: 0.00
Seq primer: -40up from Gibco
High quality sequence stop: 414.

FEATURES
source
Location/Qualifiers
1. .428
/organism="Homo sapiens"
/db.xref="taxon:9606"
/clone="IMAGE:2134523"
/clone_lib="NCI-CGAP_Lym12"
/tissue_type="Lymphoma, follicular mixed small and large cell"
/lab_host="DH10B"
/note="Organ: lymph node; Vector: PCMV-SPORT6; Site.1: SalI; Site.2: NotI; Cloned unidirectionally. Primer: 01190 dT. Average insert size 1.25 kb. Life Technologies catalog #: 11547-015"

BASE COUNT 121 a 86 c 120 g 101 t
ORIGIN

Query Match 13.4%; Score 372; DB 9; Length 428;
Best Local Similarity 99.8%; Pred. No. 1.5e-184;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2350 tcaagtaatgatagtcaggaagcagtcacaatgccagcagatagagtgagta 2409
|||||
DB 428 TCAGTATATGATGATGCGAGGAGGAGCAATGCCAGGAGATAGGGGTGGTACTG 369
|||||

QY 2410 gtagaacccaaccttaaacgtgaagacagtcgcggcctaaatccctactgaatga 2469
|||||
DB 368 GTGAACCCCAACCTTAAAGCTGAAGACAGTCCGGCTAATCCCTAATCTGATGATGAG 309
|||||

QY 2470 cctgtctcccatctgtgtgtcttcctcccgatgatacccaaccttactatttcg 2529
|||||
DB 308 CCTGCTTCCCACTTGGTGTGCTTCTCCGATGATCCCAACCTTACCTATTATTACG 249
|||||

QY 2530 tataccctccctccctaaatgttttacactgtctgtgcccacactttgagtgct 2589
|||||
DB 248 TATACCTGCGCTTCTCTAATGTTTTCATCTGCTGTGCCACCTTTTGAATGCTGCT 189
|||||

QY 2590 ttgcatacttaacaaatcagtcagtcagtcataatccctatcttgagccataaaagaccag 2649
|||||
DB 188 TTGCATACTTACAAATCAGTCAACGTATTCCTCTAATTCGAGCCCAATAAAGACCCAG 129
|||||

QY 2650 actcagctgcagtcagtcaggaagaatccaccctgcgtgtgaggttgaggacacccctcga 2709
|||||
DB 128 ACTCAGCTGCACTGAGAGAGAAATCACCTGCTGTGGGGTGGGAGCACTCCCTGCA 69
|||||

QY 2710 tccctcccaactgagagctgtctcttgcataaataatctcttaccatcc 2769
|||||
DB 68 TCCCCTTCGACTGAGAGCTGTCTTTTGCTCAATAAATCTTTTTCACCATCTCTCAG 9
|||||

QY 2770 cct 2772
|||||
DB 8 CCT 6

RESULT 27
LOCUS A1953623 458 bp mRNA linear EST 06-SEP-1999
DEFINITION wg23a03.x1 NCI-CGAP_K1d11 Homo sapiens cDNA clone IMAGE:2472076 3',
mRNA sequence.
ACCESSION A1953623
VERSION A1953623.1 GI:5745933
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS 1 (bases 1 to 458)
TITLE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 447.

FEATURES
source
1..458
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI-CGAP_Kid11"
/lab_host="DH10B"
/note="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI-CGAP_Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonids 132376-132911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 129 a 99 c 125 g 105 t
ORIGIN

Query Match 13.4%; Score 372; DB 9; Length 458;
Best Local Similarity 100.0%; Pred. No. 1.5e-184;
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2316 tctcactgtgtagtggtgtgtgtaaacccggtcaagtataatgtgagagga 2375
|||||
458 TCTCACTGTGATGGTGGTGTGGAAACCCGCTCAGTAAATGATGCGAGAGCA 399
2376 gtcaaatgccagcagatagagggtgtgtacctgtgtgaaccacacctaaga 2435
|||||
398 GTCAATGCCAGCAGATAGAGGAGTGTGCTGCGTAAACCACTTAAAGCTAAGA 339
2436 cagtcggcgctaaatcctatactgaatgtgagaacctgtctccattgtgtgttc 2495
|||||
338 CAGTCCCGGCTAAATCCTACTACTGAAATGAGAACTGCTCCATTGGTGCTTC 279
2496 ctccagtgatcccaaccttaacctattgaagtatactgccttccctaattgttt 2555
|||||
278 CTCGGATTATCCCAACCTTCACCTATTTAGTATACCTGCGCTTCTTAAATGATT 219
2556 ttacaatgtgtgacccacctttgaatgtgtgacctgtcactataaatacgaacgt 2615
218 TTACACTGCTGTGCGCACCTTTTGAATGGTGCTTTGCACTTACAAATCACTCAAGT 159
2616 gtatccctattctgagccataaagaccagactcagctgcaatgagagagaatc 2675
|||||
158 GTATCCCTAATTCGAGGCCATAAAGACCGAGACTGACGTGACATGAGAGAAATC 99
2676 accctgctgtg 2687
|||||
98 ACCCTGCTGTGG 87

RESULT 28
AM140124/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. CDNA Library Preparation: M.B. Soares lab clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html The following repetitive elements were found in this cDNA sequence: 58-184, >MER52#unknown 348-425, >MER52#unknown
Seq primer: M13 Forward
POLYA=yes.

FEATURES
source
1..501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:2719247"
/clone_lib="NCI-CGAP_Sub3"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The NCI-CGAP_Sub3 library is a subtracted library derived from the NCI-CGAP_Sub1 library, which is a subtracted library derived from B1. B1 constitutes a mixture of 21 normalized or subtracted NCI-CGAP libraries: NCI-CGAP_Co4, NCI-CGAP_Pr22, NCI-CGAP_Pr28, NCI-CGAP_Co10, NCI-CGAP_Co16, NCI-CGAP_Kid5, NCI-CGAP_Kid12, NCI-CGAP_Kid3, NCI-CGAP_Kid11, NCI-CGAP_Lym2, NCI-CGAP_Br2, NCI-CGAP_Co8, NCI-CGAP_CLI1, NCI-CGAP_Le12, NCI-CGAP_Brn23, NCI-CGAP_Lu5, NCI-CGAP_Lu24, NCI-CGAP_Lu19, NCI-CGAP_GC4, NCI-CGAP_GC6, NCI-CGAP_Brn25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:
NCI-CGAP_Kid3 pool 1 LLM 3334-3337, 3682-3683, 3798-3803 (IMAGE clonids 132376-132911, 1456008-1456775, 1500552-1502855); NCI-CGAP_Kid5 pool 1 LLM 3338-3342, 3722-3725, 3776-3778 (IMAGE clonids 1323912-1325831, 1471368-1472903, 1492104-1493253); NCI-CGAP_Lu5 pool 1 LLM 3575-3582, 3851-3854 (IMAGE clonids 1414920-1417991, 1520904-1522439); NCI-CGAP_GC4 pool 1 LLM 3164-3167, 3716-3720, 3733-3735 (IMAGE clonids 1257096-1258631, 1469064-1470983, 1475592-1476743); NCI-CGAP_Pr22 pool 1 LLM 2457-2459, 2758-2759, 3062-3068 (IMAGE clonids 985608-986759, 1101192-1101959, 1217928-1220615); NCI-CGAP_Co10 pool 1 LLM 2644-2653, 2871-2872 (IMAGE clonids 1057416-1061255, 1144584-1145351). Subtraction was performed as previously described [Bonaldo, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches to Facilitate Gene Discovery. Genome Research 6, 791-806.
TAG_LIB=NCI-CGAP_Kid5

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.lnl.gov/db/rrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 469.

FEATURES

source

1. 486
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2304013"
/clone_lib="NCI_CGAP_GC6"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA
from the normalized library NCI_CGAP_GC4 was prepared, and
ss circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clonesDB
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 136 a 108 c 133 g 109 t
ORIGIN

Query Match 12.8%; Score 355; DB 9; Length 486;
Best Local Similarity 99.8%; Pred. No. 1.4e-175;
Matches 405; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2282 agaatgctggcgcatttcagagagcccttcctcactctgtagtgggtggtg 2341
|||||
DB 486 AGAATTGCGGCATTTTCAGAGAGCCCTGCTCTCTCACTGTGATGGGTGG 427
|||||
QY 2342 aaaccggtcaggtgaatgaatgagcagagcagaatgcccagcagataggg 2401
|||||
DB 426 AAACCGGTCAGGTATGATGAGGAGGAGGCAATGCCCCAGATAGGGGG 367
|||||
QY 2402 ggtactggtgaaacccaactaaagctggaagacagtcctcggtactaatc 2461
|||||
DB 366 GGATCTGTGTGAACCACTTAAGCTGAACAGTCCGGCTTAATCTCATCTA 307
|||||
QY 2462 attgaaacctgtctccattgtgtgtcttcctccagattatcccaacctta 2521
|||||
DB 306 ATTGAACACTGTCTCCATTGGTGTGCTTCTCCATTTATCCCAACCTTACCT 247
|||||
QY 2522 atttaacgtataccgtcccttcctcactgattttacactgctgtgccac 2581
|||||
DB 246 ATTTTACGTATACCTGCTCCCTTCCCTTAATGTTTAACTGCTGGGCCACCTTTGAG 187
|||||
QY 2582 tgggtccttgcatactatcaaatcagtcgaagtgtatccctattctgagcc 2641
|||||
DB 186 TGGTGCTTTGATCTTCAATCAATCAGTCAACCTGTATTCCTCATTTGAGCCCAATAA 127
|||||
QY 2642 agaaccaagctcagctcagtgagagaaatcacctcgtcgtg 2687
|||||
DB 126 AAGACCCAGACTCAGCTGCAAGAGAGAGAAATCACCTGCTGTGG 81
|||||

RESULT 31

BF476173/c 473 bp mRNA linear EST 05-DEC-2000
LOCUS naa29c11.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3257852
DEFINITION 3', mRNA sequence.

ACCESSION BF476173
VERSION BF476173.1 GI:11546989
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 473)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).

JOURNAL Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

COMMENT

Tissue procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.lnl.gov

Seq primer: -40UP from Gibco
High quality sequence stop: 464.

FEATURES

source

1. 473
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3257852"
/clone_lib="NCI_CGAP_Kid11"
/lab_host="DH10B"
/note="Organ: kidney; Vector: p773D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(clonesDB 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo.

BASE COUNT 133 a 104 c 126 g 110 t
ORIGIN

Query Match 12.7%; Score 351; DB 10; Length 473;
Best Local Similarity 99.6%; Pred. No. 1.8e-175;
Matches 451; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2310 tgccttcctcactctggtgtggtgtgtggaacccggtcaggtgaatgtagtggcag 2369
|||||
DB 458 TGCCTTCCTCACTGTGATGAGTGGGTGTGGAACCCGCTCAGTAAATGATAGTGGCAG 399
|||||
QY 2370 gaggcagtaaatgcccagcagataggggtgtgtactcgtgtgaaacccaactaaagc 2429
|||||
DB 398 GAGGCGATGTAATGCCACAGAGATAGGGGTGATACCTGTGAACCCCAACCTTAAGC 339
|||||
QY 2430 tgaagcagctcccggttaactcactgaattgagaacctgtcttcctcattgtgt 2489
|||||
DB 338 TGAAGCACTCCCGCTTAATCTCTACTGAATGAGAACCTGTCTCCATTGTGGTGT 279
|||||
QY 2490 gcttcctcagattatcccaaccttcacattttacgtatcactcgtcccttcctaat 2549
|||||
DB 278 GCTTTCCTCGAATGATGCCCAACCTTCACATTTTACGTATACCTGCTTCTCTAAT 219
|||||
QY 2550 tgggtttacactgctgtgccacactttgagtggtgcttgcatacttacaatcag 2609
|||||
DB 218 TGGTTTTACACTGTGTGTCGCCACCTTTGAGTGTGGCTTGCATTTACAAATCAAT 159
|||||
QY 2610 caacgtgatccctcattctgaagcccaaaaagcccgactcagctgagtgagagga 2669
|||||
DB 158 CAACGTGATTTCCCTATTCTGAGCCCAATAAAGACCCAGACTGACGTGCAATGAGAGGA 99
|||||
QY 2670 gaatatccctcgtctggaagcttggaagcaactccctgcatccctcctcactgaagct 2729
|||||
DB 98 GAATTCACCTCTGTGGGGGTTGGGAGACACTCCCTGATCCCTCTCCATGAGAGCT 39
|||||
QY 2730 gtctcttgcataaataatcttcttaccga 2762
|||||
DB 38 GTTCTTTCCTCAATTAATTTCTTTCTACCA 6
|||||

RESULT 32
AA218658 385 bp mRNA linear EST 12-MAR-1998
LOCUS z936904.f1 Stratagene NT2 neuronal precursor 937230 Homo sapiens
DEFINITION cDNA clone IMAGE:649878 5' similar to gb:M29580 ZINC FINGER PROTEIN
7 (HUMAN);, mRNA sequence.
ACCESSION AA218658
VERSION AA218658.1 GI:1832766
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 385)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,
J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wyllie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert length: 2813 Std Error: 0.00
Seq primer: -28ml3 rev1 ET from Amersham.

FEATURES
source
1..385
/organism="Homo sapiens"
/db_xref="GDB:5276467"
/db_xref="taxon:9606"
/clone="IMAGE:649878"
/clone_lib="Stratagene NT2 neuronal precursor 937230"
/tissue_type="neuroepithelial cells"
/dev_stage="Ntera-2 neuroepithelial cells"
/note="Organ: brain; Vector: plasmid; Site: 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Uninduced, exponentially growing neuroepithelial
cells (Ntera-2/cl.DJ). Average insert size: 1.0 kb.
Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGCGACGAG
3' -3' adaptor sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3'."

BASE COUNT 89 a 102 c 101 g 93 t
ORIGIN

Query Match 12.6%; Score 350; DB 9; Length 385;
Best Local Similarity 100.0%; Pred. No. 5.9e-173;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 207 ggcgtcggcattcaggagatgctgtgtacccacccagagagatggtgtg 266
DB 36 GCGTTCGTGGCAATTCAGGATGTGGCTGTGACTTCACCCAGAGAGAGGAGTTGTG 95
QY 267 agccctgtcagagaccctgcacagggagtgatctgtgaaactatatacatctgtc 326
DB 96 AGCCTGTGCAGAGAGACCTGCACAGGGAGTGATGCTGAGACTTATACCATCTGTC 155
QY 327 tcaatggaatcattcttctaaccacaaatcattgtctcagctggaagcgaggaagcg 386
DB 156 TCACGGAATTCATCTTAAACCAAACTATTGCTCACCTGAGCGAGCGAGGACG 215
QY 387 cccttgaagagagagaaatgttcactggaacctctgtcgaagatcgagacgaat 446
DB 216 CCTCTGAGAGAGAGAAATGTCACCTGTGTCGCAATCGAAGCCAGAAAT 275
QY 447 caacttaagccctcctcctctgatttctccagtcagcaagcttcagccaacatgtg 506

DB 276 CAACTAGTCCCTCTGCTCTGTGATTTCTCCAGTCAAGACGCTCAGCCACATGTG 335
QY 507 tggcgtgacatctctcaccagctgtttcagaattatggaagaacc 556
DB 336 TGGCTGAGTCAATCTCTCAGCTGTGTTCAAGATTATGGAGGAATTC 385

RESULT 33
A1986444/c 500 bp mRNA linear EST 20-OCT-2000
LOCUS w265d09.x1 NCI CGAP_Mel15 Homo sapiens cDNA clone IMAGE:2562929 3',
DEFINITION mRNA sequence.
ACCESSION A1986444
VERSION A1986444.1 GI:5813721
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 500)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Chris Moskalk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: Life
Technologies, Inc. CDNA Library Arrayed by: Christa Prange, The
I.M.A.G.E. Consortium DNA Sequencing by: Washington University
Genome Sequencing Center
Gene distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/db/rip/image/image.html
Insert length: 972 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 403.

FEATURES
source
1..500
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2562929"
/clone_lib="NCI CGAP_Mel15"
/tissue_type="malignant melanoma, metastatic to lymph
node"
/lab_host="DH10B"
/note="Organ: skin; Vector: PCMV-SPORT6; Site: 1:
SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies."

BASE COUNT 139 a 105 c 132 g 124 t
ORIGIN

Query Match 12.6%; Score 350; DB 9; Length 500;
Best Local Similarity 99.8%; Pred. No. 6.2e-173;
Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2287 tgcctgctcatttcagagaccctcctcctcactgtgattggtgtgaaacc 2346
DB 500 TCTGCTCTCATTTTCAAGAGCCCTCCCTCTCCTCACTGTGATGGTGTGGAATTC 441
QY 2347 cggtagagtaatgtagtgagagagagagagagagagagagagagagagagagag 2406
DB 440 CGGTCAAGTAATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 381
QY 2407 ctggtgaaccacaccltaagaactgaagacagtcctcgcttaaatcctcactgaatga 2466
DB 380 CTTGGTGAATCCACACTTAAGCTGAAGACAGTCCCGCTTAATCCTCACTGAATGA 321
QY 2467 gaacctgtctcccatctgtgtgtcttctcccgatgtatcccaaccttaaccttt 2526
DB 320 GAACGTGTCCTCCATTTGGTGTGCTTCCTCCGATTTGATTCACCAACCTTCATTTT 261

QY 2527 acgtataccctcccttcccaattgttttacaactgctgcccacctttagtgg 2586
|||||
Db 260 AGCTATACCGCCCTTCCATTTGTTTACAGCTGCTGGCCACTTTTACGAGTG 201
QY 2587 ccttgcatactacaataagtcacgltgaltccctcattcttgagccataaagacc 2646
CcTTTGCATCTTCAATCAAAATGATCAACAGTGTATTTCCCTATTTTGAGCCATTAAGAGACC 141
Db 200
QY 2647 caactcagctgagtgagagagaataacccctgctgag 2687
|||||
Db 140 CAGACTCAGCTGAGTGAGAGAAATCACCTGCTGTGG 100

RESULT 34
LOCUS BE395838 551 bp mRNA linear EST 21-JUL-2000
DEFINITION 60131007F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3631603 5',
mRNA sequence.
ACCESSION BE395838
VERSION BE395838
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC <http://img.nci.nih.gov/>.
1 (bases 1 to 551)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LCM321 row: j column: 20
High quality sequence stop: 545.

FEATURES
source
1..551
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3631603"
/clone_1lb="NIH_MGC_44"
/tissue_type="endometrium, adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(C). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 158 a 129 c 153 g 111 t
ORIGIN

Query Match 12.4%; Score 345; DB 10; Length 551;
Best Local Similarity 99.4%; Pred. No. 2.7e-170;
Matches 495; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1217 cctgaagtaaacctcattaccacacagagggcgacactgggagagacttatgttg 1276
|||||
Db 34 CCGAAGTCAAACTTATACCCACAGAGGGCGACACTGGGAGAGAGCCCTTATGTGG 93
QY 1277 caggagatctggcgctgttgcagacattacacccctgttagacacagaagacaca 1336
|||||
Db 94 CAGGGATGTGGGCGTGGCTTTCGCCAGCATTCACACCTGGTGTAGACACAGAGACACA 153
QY 1337 ttcagagagaagccttacctgtaggaggtgtgagcaagcttagcagaagtcaca 1396

Db 154 TTGAGAGAGAAACCTTACTTTCAGGAGGTGTGAGCAAGCTTGAGCCAGAAATGCACA 213
QY 1397 cctcatcagacacttaagagacacacacagagagaagacttatgtatgcagagaatg 1456
Db 214 CTTTCATCAGACACTTAAAGGACACACAGAGAGAGAGCCCTTATGTGTGACACAAATGTGG 273
QY 1457 ggcacacttaagcttggaatcaaacctcaaaaacacacacagagagacactcaagggttaa 1516
Db 274 GCCTCACTTAGCTGGAATTCAAACCTCAAAACACACAGAGACACTCAGGGGTAA 333
QY 1517 acctatgtctgccttgagtgcgggcagltgctttagcctgaagtcacacttaacaaaca 1576
Db 334 ACCTTATGTCTGCTGAGTGCAGTGCAGTCTTACCTGAAGTCTAACCTTAAACAAACA 393
QY 1577 ccgagagtcacacacaggggagagaagccattgtatgtagaggtgtggtcgaggtctac 1636
|||||
Db 394 CCGAAGGTCAACACCGGGGAGAGCAATTTGTATGTAGGAGTGTGGCGAGCTTTAGC 453
QY 1637 ccggaatcaaccctgatcacgaccagagagacacactcagggagagaagccattgtatg 1696
Db 454 CCGGAATCAACCCCTGAGCAGCAGCAGCAGACACACTCAGGGAGAACCCATTGTATG 513
QY 1697 tgctgagtgtagagag 1714
Db 514 TCCTGAGTGTGAGCAGAG 531

RESULT 35
LOCUS A1860707 508 bp mRNA linear EST 07-MAR-2000
DEFINITION w151d12.x1 NCI-CGAP_Ut1 Homo sapiens cDNA clone IMAGE:2424983 3',
mRNA sequence.
ACCESSION A1860707
VERSION A1860707
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
1 (bases 1 to 508)
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Christopher Moskajuk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 1940 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 406.

FEATURES
source
1..508
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2424983"
/clone_1lb="NCI-CGAP_Ut1"
/tissue_type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 Kb. Life Technologies catalog #:
11538-014"

BASE COUNT 140 a 110 c 138 g 120 t
ORIGIN

Query Match 12.4%; Score 344; DB 9; Length 508;
Best Local Similarity 99.4%; Pred. No. 9, 1e-170;
Matches 494; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2276 tgaagagaatgcgcgcgcatttcgaagagccgcgccttcctcactctgagtgatg 2335
DB 508 tgaagagaatgcgcgcgcatttcgaagagccgcgccttcctcactctgagtgatg 449

QY 2336 ttgtygaaacccgcgtcaggtlaatgtagtgcaggaagagcagtaacatgcagcagata 2395
DB 448 ttgtggaacccgcgtcaggtlaatgtagtgcaggaagagcagtaacatgcagcagata 389

QY 2396 ggggtgggtacctggttgaaaccccaactaaagctgaagacagctcccggttaactccta 2455
DB 388 ggggtgggtacctggttgaaaccccaactaaagctgaagacagctcccggttaactccta 329

QY 2456 tactgaatgaacactgctcctccattgtgtgtcttcctccgaattatcccaacct 2515
DB 328 tactgaatgaacactgctcctccattgtgtgtcttcctccgaattatcccaacct 269

QY 2516 tcacctatttaacgtataccgtcccttcctcaattggttttaactgctgtgcccacct 2575
DB 268 tcacctattttacgtataccgtcccttcctcaattggttttaactgctgtgcccacct 209

QY 2576 ttgagtggtgctcttgatctactcaaatcagtgtaacgtatctccctatcttgagcc 2635
DB 208 ttgagtggtgctcttgatctactcaaatcagtgtaacgtatctccctatcttgagcc 149

QY 2636 cataaagacccagactcagctcagctgaggaagagaatcacctgctgtgaggtgag 2695
DB 148 cataaagacccagactcagctcagctgaggaagagaatcacctgctgtgaggtgag 89

QY 2636 gacacccctcgtcgtccctctcctcagctgaggtgctcttcctgccaataaattcttt 2755
DB 88 gacacccctcgtcgtccctctcctcagctgaggtgctcttcctgccaataaattcttt 29

QY 2756 ctaccatctcctcact 2772
DB 28 ctaccatctcctcact 12

RESULT 36
AI289399/c 399 bp mRNA linear EST 21-DEC-1998
LOCUS qw32e05.x1 NCI_CGAP_Ut4 Homo sapiens cDNA clone IMAGE:1992800 3',
DEFINITION mRNA sequence.
ACCESSION AI289399
VERSION AI289399.1 GI:3931430
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 399)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 2932 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 395.

FEATURES
Source Location/Qualifiers
1..399
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1992800"
/clone_11b="NCI_CGAP_Ut4"
/tissue_type="serous papillary carcinoma, high grade, 2
pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-Sport6; Site: 1; Salt;
Site: 2; NCI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.48 kb. Life Technologies catalog #:
11542-016"

BASE COUNT 114 a 76 c 113 g 96 t
ORIGIN

Query Match 12.4%; Score 343; DB 9; Length 399;
Best Local Similarity 99.7%; Pred. No. 2, 9e-169;
Matches 393; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2375 agtcaaatgcccagcagataggggttggtacctgtgtgaacccaacttaagctgaag 2434
DB 399 agtcaaatgcccagcagataggggttggtacctgtgtgaacccaacttaagctgaag 340

QY 2435 acagtcggcgctaaatcctcatatgaattggaacctgtctccattgtgtgcttt 2494
DB 339 acagtcggcgctaaatcctcatatgaattggaacctgtctccattgtgtgcttt 280

QY 2495 cctccgattgatacccaaccttaactatcttaacgtatataccgaccttcctaattgct 2554
DB 279 cctccgattgatacccaaccttaactatcttaacgtatataccgaccttcctaattgct 220

QY 2555 tttaacgtgctgtgcccaccttttgagtggtgctttgcatattcaataatgaatgaag 2614
DB 219 tttaacgtgctgtgcccaccttttgagtggtgctttgcatattcaataatgaatgaag 160

QY 2615 tttatcccttattctgagcccaataaagacccaagactcagctgagtgaggaagaat 2674
DB 159 tttatcccttattctgagcccaataaagacccaagactcagctgagtgaggaagaat 100

QY 2675 caacctgctgtggaagttgaggacacatccctgcagatccctctcactgagagctgtct 2734
DB 99 caacctgctgtggaagttgaggacacatccctgcagatccctctcactgagagctgtct 40

QY 2735 ttgtcccaataaattcttttaccatccctca 2768
DB 39 ttgtcccaataaattcttttaccatccctca 6

RESULT 37
AA218659/c 392 bp mRNA linear EST 12-MAR-1998
LOCUS zq96g04.s1 Stratagene NT2 neuronal precursor 937230 Homo sapiens
DEFINITION cDNA clone IMAGE:649878 3', mRNA sequence.
ACCESSION AA218659
VERSION AA218659.1 GI:1832767
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 392)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Gelsel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin,
J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wylie, T., Waterston, R. and Wilson, R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800

Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LINT; contact the
IMaGE Consortium (info@image.lln.gov) for further information.
Insert length: 2813 Std Error: 0.000
Seq primer: -41m13 fwd. ETT from Amersham.

FEATURES	Location/Qualifiers
source	1. .392

BASE COUNT	111 a	82 c	108 g	91 t
ORIGIN				

Query Match	12.3%;	Score 340;	DB 9;	Length 392;
-------------	--------	------------	-------	-------------

Matches 390; Conservative 0; Mismatches 1; Indels 0; Caps 0;

Db 391 TGGCAGGAGCAGTCAATGCCCAGGCAGATAGGGTGGGTACCTGGTGAACCCCAACT 332

QY . 2424 taaagctgaagacagtcgccgctaatactcctatactgaattgagaacctgtcttccatt 248:

Db 331 TAAAGCTGAGACAGTCCCGGCTAAATCCTCATCTGAATTGAGAACCTGTCTTCCCATTT 27

OV 2484 t a g t a t c t t t c c t c c q a t t g a t c c c a a c c c t t c a c c t a t t t t a c g t a t a c c t a c c c t t 2541

Db 271 TGGTGTCTTTCCTCCGATTGATCCCAACCTTACCTATTTCACGTATACCTGCCCTT 212

ov 2544 cctaattgattttacactgctatgccaccccttttaaatgataacctttacatacttacaa 2601

dh 211 CCTATTGGTATTACACCTGGCTGACCTATTGACATGCTTACAA 152

Ov 350A atcagttcaaaacattat+ccccctatat+crraaccccaa+aaaaaaaaccaccaaatcacacn+rcaactr 366

[illegible][illegible]

2

[illegible]

RESULT	38
LOCUS	AJ361705/c
DEFINITION	AJ361705 395 bp mRNA linear EST 16-FEB-1999
ACCESSION	gq18f04.x1 NCI-CGAP-CLL1 Homo sapiens cDNA clone IMAGE:202187 3',
VERSION	AJ361705 mRNA sequence.
KEYWORDS	AJ361705.1 GI:4113326
SOURCE	EST.
ORGANISM	human.
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 395)
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .

TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP)
JOURNAL	Tumor Gene Index Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D.

Email: cgap@lmln1.llnl.gov
Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever, M.D., Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bdrrp/image/image.html
Insert length: 1512 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 385.

BASE COUNT	113 a	76 c	112 g	94 t
------------	-------	------	-------	------

Query Match	12.38;	Score 340;	DB 9;	Length 395;
-------------	--------	------------	-------	-------------

```

Best Local Similarity 99.76; Pred. NO. 1.1e-107;
Matches 390; Conservative 0; Mismatches 1; Indels 0; Gaps

```

2374 caatcaaatgcccaqccaataaqqqtacctaqtgaaaaccaaccttaaaqctgaa 2433

Db 395 CAGTCAATGCCCAGGCAGTAGGGTGGTACCTGGTGAACCCACCTTAAGCTGA 336

2434 gaccatcccccaatacctcctatactgaaattgaagaaacctatctcccaattgatatnctt 2497

335 CACCCTCCCGCTAATCTGACGACCTGTCCTCCATTGCTGCTCTT 276

On 3104 +cctccccat+catccccccccct+caacatattt+aaacat+at+acccct+ccccct+ccct+aaatt+ 3555

[illegible]

.....

[illegible]

22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051

RESULT	39				
AW958936					
LOCUS					
AW958936					
	513 bp	-mrna	linear	EST	01-JUN-2000

DEFINITION EST3711006 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.
ACCESSION AM958936
VERSION AM958936.1 GI:8148620
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 513)
Hedde, P., Qi, R., Abernathy, K., Dharap, S., Gasparé, R., Gay, C., Holt,
I. E., Saeed, A. I., Sharov, V., Lee, N. H., Yeatman, T. J. and
Quackenbush, J.
TITLE Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tifgr.org
Plate: 127
Seq primer: Reverse.
FEATURES
source Location/Qualifiers
1..513
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGE"
/note="Vector: pBluescriptSKm"
BASE COUNT 130 a 142 c 102 g 139 t
ORIGIN
Query Match 12.2%; Score 337; DB 9; Length 513;
Best Local Similarity 99.7%; Pred. No. 4.5e-166;
Matches 387; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2377 tcaaatgcccagcagataggggtggtggtacccgtgtaaacccaacttaagctgaagac 2436
1 TCAATGCCAGCAGCATAGGGGTGGGTACCTGCTGAACCCCACTTAAGCTGAAGAC 60
QY 2437 agtcccgagtaaatccctcaactgaatgagaacctgtctcccatgtgtgtctcc 2496
61 AGTCCCGAGTAATCCCTCACTGAATGAGAACCTGCTTCCATTTGGTGTCTTCC 120
QY 2497 tccgattatcccaacctcaactatttaagtaactgcgccttctctaattgtttt 2556
121 TCCGATTGATCCCAACCTTCACTATTTTACGTATACCTGCCCTTCTTAATTGGTTT 180
QY 2557 taactgtgtgcccacctttagtgtgtgcttgcataactacaatcagtcacgtg 2616
181 TAACTGCTGTGCCACCTTTGAGTGTGCTTGTGCACTTACAAATCAGTCAACGTG 240
QY 2617 tattcccatctctgagccataaagaagaccagactcagctcagtgagagagaataca 2676
241 TATTCCTCCATTTCTGAGCCCATAAAGACCCAGACTCAGCTCAGAGAGAGAAATCA 300
QY 2677 cctctgtgtgaggtgtgggagaccacctcctgcatccctctcactggaagctgttctt 2736
301 CCTGCTGTGGGGGTGGGAGCCACTCCCTGCATCCCTCCACAGAGAGCTGTTCTTT 360
QY 2737 tgcataaatactcttctaccatc 2764
361 TGCTCAATAAATCTTTCTTACCCATC 388
RESULT 40
BI256434 794 bp mRNA linear EST 17-JUL-2001
LOCUS 602974454F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5113861-5',
DEFINITION mRNA sequence.
ACCESSION BI256434
VERSION BI256434.1 GI:14810838

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 794)
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@ds-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/MLN at:
<http://image.llnl.gov>
Plate: LLAM11276 row: k column: 14
High quality sequence stop: 791.
FEATURES
source Location/Qualifiers
1..794
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:5113861"
/clone_lib="NIH_MGC_12"
/issue_type="Cervical carcinoma cell line"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.4 kb. Library prepared by Life
Technologies."
BASE COUNT 231 a 209 c 188 g 166 t
ORIGIN
Query Match 12.0%; Score 333; DB 10; Length 794;
Best Local Similarity 99.8%; Pred. No. 6.3e-164;
Matches 453; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 333 gaatcattctcttaaaccaactcattgtctcagctgtagagcgagggagcgccctg 392
189 GAATTCATCTCTTAACCAAACTCATGCTCAGCTGGACCGAGGCGGCCCTGTG 248
QY 393 agagaagaagaagaatgctcactgagacctctgtccaagatcgaaagcaaaatcaact 452
249 AGAGAGAGAGAAATGTCACCTGGACCTGTGTCCAGAAATCGAAGCCAGAAATTCATT 308
QY 453 agtccctcctgcctctgatttctcagctcagcaagctctcagccaatgtgtgctg 512
309 AGTCCCTCGCTGCGCTGTGATTTTCTCAGTCAAGCAAGCTCTCAGCAACATGTGGTG 368
QY 513 agtcatctctcagctgttttcaagttatgggcaagaatctctcactcctgggaaa 572
369 AGTCATCTCTCAGCTGTTTTCMACTTATGGCGAGAAATCTCTCCACCTGGGAAAA 428
QY 573 caatacagaagaataagaacaagaagcagagatccattctgtcttagtggaaagaaga 632
429 CACTATCCAGAAATCAGAAACACAGACAGATCCATTGCTTGTGTGCAAGAGAGAA 488
QY 633 tgaattcaagaagga-gaagactcagactcctgttttgagagagtaagaataatgacac 691
489 TGGAATTCAGAGAGAGAGAGAGAGACTCCAGACTCGTGTGGAGAGGTAAAGCAAAATG 548
QY 692 ttcaaaagcaattccagccacactgaaagaacaagaagcagagcagttccaaggaagaca 751
549 TTCAAAGGCACTTCCAGGCCACCTGAAGAACACAGCAGCAGCAGTCCAAAGAGACAA 608
QY 752 cacagctgtgtatagaggtccagccctgaagc 785
609 CACAGTGTGTGATATAGGGTCCAGCCCTGACCG 642

[illegible]

	Db	143	TATTGTGAGCCCATATAAAGACCCACACTCAGCTGCAGTGAGGAGAAGAAATCACCTGCTG	84
Oy	2685	tggaggttggggaaccactccctcgtacatcccctcttcacctgaagagtgtctcttgccaat	2744	
Db	83	TGGGGGTTGGGGACACAGTCCCTGCATGCCCTTCACATGAGAGCGTGTCTTGGTCATAT	24	
Oy	2745	aaaattctttctaccatctc	2767	
Db	23	AAAATTCCTTCTACCACATCCTC	1	
RESULT	42			
LOCUS	AA484745/c	378 bp	mRNA	linear EST 19-AUG-1997
DEFINITION	nebl1h06.s1 NCI_CGAP_Ew1 Homo sapiens cDNA clone IMAGE:910715, mRNA sequence.			
ACCESSION	AA484745			
VERSION	AA484745.1	GI:2214130		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
TITLE	1 (bases 1 to 378)			
JOURNAL	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.			
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997)			
	Contact: Robert Strausberg, Ph.D.			
	Email: cgaps-remail.nih.gov			
	Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D., Ph.D.			
	cDNA Library Preparation: David B. Krizman, Ph.D.			
	cDNA Library Arrayed by: Greg Lennon, Ph.D.			
	DNA Sequencing by: Washington University Genome Sequencing Center			
	clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:			
	www-bio.lnl.gov/bdrr/image/image.html			
	Insert Length: 457 Std Error: 0.00			
	Seq primer: -41ml3 fwd. ET from Amer sham			
	High quality sequence stop: 377.			
FEATURES	Location/Qualifiers			
source	1..378			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	* /clone="IMAGE:910715"			
	/clone_lib="NCI CGAP_Ew1"			
	/tissue_type="Ewing's sarcoma"			
	/lab_host="DH10B"			
	/note="Vector: pAMP10; mRNA made from Ewing's sarcoma, cDNA made by oligo-dT priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."			
BASE COUNT	111 a 67 c 108 g 92 t			
ORIGIN				
Query Match	11.5%: Score 318: DB 9: Length 378:			
Best Local Similarity	99.7%: Pred. No. 4.5e-156:			
Matches 368: Conservative	0: Mismatches 1: Indels 0: Gaps 0:			
Oy	2400	tgggtacctgtgtaaaccaacctaaagaacgtaagacagtcgcgctaataatccatacat	2459	
Db	378	TGGGTAAGTGGTGAAGCAACCACTTAAGCGTAGAACACAGTCCGGCTMAATCCTATACT	319	
Oy	2460	gaattgagaacctgtctctccatttgtgtgctcttcctcgatytgatcccaacctcac	2519	
Db	318	GAAATGACAACCTGCTCTTCCATTTGGTGGTCTTCTCCGATGTGATCCCAACCTTCAC	259	
Oy	2520	ctatttacgatatacctgtcccttctccaataggittttaacctgtctgcccaccttt	2579	

Db	258	CATATTTCAGATACCTGCCCTTTTCCTAATTGGTTTTTACAAGTGCTGCGCCACTTTTG	199
OY	2580	agtgtgcttcttgacttaataacagtcgaagcgtgatccctattcttgcagcca	2639
Db	198	AAGTGTGCCTTTGGCATCTTAAATAATCAGTCACACTGTATTCCTCTATTCGAGCCATA	139
OY	2640	aagaagcccaagactgagcgtagtgagagaaatcacccctgcctcttgtagtctggagcc	2699
Db	138	AAAGACCAGACTCAAGCTGCACTGAGGAGAGAATAACCCCTGCTGTGGGGTGGAGCC	79
OY	2700	acctccctgcatccccctcccacactgagagactctttcttctgcataaaaattcttttac	2759
Db	78	ACTCCCTGCAGTCCCCTCCCACTGAGACTGTTCTTTTGGCTCAATAAATCTTTTCATC	19
OY	2760	cocatccca	2768
Db	18	CCATCTCA	10
RESULT	43		
LOCUS	A1159962/c	369 bp	mRNA linear EST 26-Oct-1998
DEFINITION	A1159962	gb08f10.x1 Soares_fetal_heart.NbHH19W Homo sapiens cDNA clone	
ACCESSION	A1159962		
VERSION	A1159962.1	GI:3693342	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE	1 (bases 1 to 369)		
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
JOURNAL	Tumor Gene Index		
COMMENT	Unpublished (1997)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs@remail.nih.gov		
	This clone is available royalty-free through LINT ; contact the		
	IMAGE Consortium (info@image.llnl.gov) for further information.		
	Insert Length: 655 Std Error: 0.00		
	Seq primer: -40m13 fwd. RT from Amersham		
FEATURES	High quality sequence stop: 364.		
SOURCE	Location/Qualifiers		
	1..369		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_image="1706443"		
	/clone_lib="Soares_fetal_heart_NbHH19W"		
	/sex="unknown"		
	/dev_stage="19 weeks"		
	/lab_host="DH10B (ampicillin resistant)"		
	/note="Organ: heart; Vector: pRT73D (Pharmacia) with a		
	modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st		
	strand cDNA was primed with a Not I - oligo(dT) primer [5'		
	TGTTACCAATCTGAATGAGGAGCGGCCCATCTTTTCTTTTCTTTTCTT 3'],		
	double-stranded cDNA was size selected, ligated to Eco RI		
	adapters (Pharmacia), digested with Not I and cloned into		
	the Not I and Eco RI sites of a modified pRT73 vector		
	(Pharmacia). Library went through one round of		
	normalization to a Cot = 5. Library constructed by		
	M.Felima Bonaldo. This library was constructed from the		
	same fetus as the fetal lung library, Soares fetal lung		
	NbHH19W."		
BASE COUNT	109 a	70 c	105 g 85 t
ORIGIN			
Query Match	11.4%; Score 315; DB 9; Length 369;		
Best Local Similarity	99.7%; Pred. No. 1.7e-154;		
Matches	365; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
OY	2397	gggtgggtacctggtgtaaacaccacttaaatgaagaagtcgccgctaactccat	2456

[illegible]

[illegible]

```

FEATURES
SOURCE
1..501
location/qualifiers
/organism="Homo sapiens"
/db_xref="GDB:565059"
/db_xref="taxon:9606"
/clone="IMAGE:152809"
/clone_lib="Soares breast 2NbHbst"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: breast; Vector: pT73D (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer (5' TGTTACCAATCTGAAGTGGGACGCGCCGCCCTTTTCTTTTCTTTTCTT 3'), double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot - 230. Library constructed by Bento Soares and M.Fatima Bonaldi."
```

Query Match	11.3%	Score 312	DB 10	Length 501
Best Local Similarity	100.0%	Pred. No. 6.9e-153		
Matches 312	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 2078	gaggacacactcagatagaacttattatgtlatatgggaatgtgttataagccattagcca	2137		
Db 1	gaggaacacactcagatagaacttattatgtlatatgggaatgtgttataagccattagcca	60		
QY 2138	gagatcacacttcatacaccacagaggaacacacacagctgtctgtcttttaagccatt	2197		
Db 61	ggagtcataacttccatcagacacacagaggaacacacacagctgtctgtcttttaagccatt	120		
QY 2198	gctatagataccaagtgtgagacattctgtgtgtatattatgcatgagactgtacgtataga	2257		
Db 121	gctatagatmccaaagtgtgagacattctgtgtgtatattatgcatgagactgtacgtataga	180		
QY 2258	ctttatcttcataccacttgaagaagaattctctgtatcttcttaaggagccctgccttc	2317		
Db 181	ctttgtatcttcataccacttgaagaagaattctctgtatcttcttaaggagccctgccttc	240		
QY 2318	cttcactgtgatactgtgtgtgtgtgtggaaccggttcagtaatgatatgtgcagagagcagt	2377		
Db 241	cttcactgtgtgatatgtgtgtgtgtgtggaaccggttcagtaatgatatgtgcagagagcagt	300		
QY 2378	caaatgtcccaag 2389			
Db 301	caaatgtcccaag 312			
RESULT 46				
LOCUS R83624				
DEFINITION	R83624 427 bp mRNA linear EST 04-AUG-1995			
	yl16h01.r1 Soares breast 3nbhst Homo sapiens cDNA clone			
	IMAGE:187633 5' similar to SP:KRI_HUMAN P10072 HKRI PROTEIN ; mRNA			
	sequence.			
ACCESSION	R83624			
VERSION	R83624.1 GI:928501			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			

REFERENCE 1 (bases 1 to 427)
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Riffe, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston, R., Williamson, A., Wohldmann, P. and Wilson, R.
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 769
High quality sequence stops: 237
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 769 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 237.
Location/Qualifiers
1. 427
/organism="Homo sapiens"
/db_xref="GDB:3818529"
/db_xref="taxon:9606"
/clone="IMAGE:187633"
/clone_lib="Soares breast 3MDHBT"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: breast; Vector: pT7D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCATCTGAAGTGGAGCGCCCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7D3 vector (Pharmacia). Library went through one round of normalization to a Col-20. Library constructed by Benito Soares and M. Fatima Bonaldo."

BASE COUNT 109 a 90 c 112 g 113 t 3 others
ORIGIN

Query Match 11.2%; Score 311; DB 10; Length 427;
Best Local Similarity 100.0%; Pred. No. 2.3e-152;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2079 agcacacacacagatagaacttctgtctatagggaatgtgtacagccttagccag 2138
|||||
Db 1 AGGACACACTCAGATAGAACTTTATGTGTATAGGGAATGTGTACAGCTTTACCCAG 60
QY 2139 gagtcatactcaacagacacagagacacacacagctgtgtgtcttttcagccatg 2198
|||||
Db 61 GAGTCATCTTATCATGACACACAGACACACAGTCTGTGCTTTTTCAGCCATTG 120
QY 2199 ctgagatacaaaagtggagcattctgtgtgtatgatgatgatgatgtactgttaagc 2258
|||||
Db 121 CTAGATACCAAGTGGAGCACTTCTGTGTGTATGATGACTGTACTGTAAAGAC 180
QY 2259 ttgtatcctcatccacctaagaagatgtgtgtcattttcaggaagccctcccttc 2318
|||||
Db 181 TTGTATCTCCATCCACCTGAAGAGAAATGTGTGCTCATTTTCAGAGCCCTCCCTTC 240
QY 2319 tcaactgtgagtgtgtgtgtgaaacccggtcaggttaagtgtgcaagagagcagtc 2378
|||||
Db 241 TCACCTGTGATGGTGTGTGAAACCGGTGATGATGATGATGATGATGATGATGATG 300
QY 2379 aaatgccagg 2389
|||||
Db 301 AAATGCCAGG 311

RESULT 47
AW513538/c
LOCUS AW513538
DEFINITION x086b03.x1 NCI_CGAP_Col17 Homo sapiens cDNA clone IMAGE:2737709 3',
mRNA sequence.
ACCESSION AW513538
VERSION AW513538.1 GI:7151616
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 410)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life Technologies
, Inc. cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E.
Consortium DNA Sequencing by: Washington University Genome
Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
image.lnl.gov/image/html/lresources.shtml

FEATURES
source
1. 410
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2737709"
/clone_lib="NCI CGAP Col17"
/tissue_type="juvenile granulosa tumor"
/lab_host="DH10B"
/note="Organ: Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies."

BASE COUNT 116 a 86 c 111 g 96 t 1 others
ORIGIN

Query Match 11.1%; Score 308; DB 9; Length 410;
Best Local Similarity 99.5%; Pred. No. 8.6e-151;
Matches 408; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2352 aggtatgtatgtgtggaagagcagtcacaaatgcccaagagatgaaggtgtgtactgt 2411
|||||
Db 410 AGGTAATGATAGTGGAGGAGGAGCAATGCGCAGGAGATGAGGAGGATCCTGGT 351
QY 2412 gaacccaacacctaagcgtgaagacagtcocggcctaactccatactgaatgaagac 2471
|||||
Db 350 GAACCCAACTTAAGGCTGAACAGACAGTCCGGCTTAATCTCATCTGATGAAAC 291
QY 2472 tgcctccaatgtgtgtgtcttcctccgattgatcccaacccctcaactatttcgta 2531
|||||
Db 290 TGTCTTCCCAATTTGGTGTGCTTCTCCGATGATGATCCCAACCTTCACCTATTTCGTA 231
QY 2532 taccctcccttccttaattgttttaacagctgtgtgtgtgtgtgtgtgtgtgtgt 2591
|||||
Db 230 TACCTCCCTTCTCTTAATGCTTTTACACTGCTGTGCTCCACCTTTTGTGCTGCTTT 171
QY 2592 gcaacttaacaatcagtcacagtgatctcccatctctagcaccataaagaagccagac 2651
|||||
Db 170 GCATCTTACAAATCAGTCAAGCTGATATCCCTATTTCTGAGCCATAAAGACCCAGAC 111
QY 2652 tcagctgcagtgagagagaatcacccctgcgtgtgtgtgtgtgtgtgtgtgtgtgt 2711
|||||

Db 110 TCAGNTCAGTGAAGAGAGAAATCCCTGCTGTGGGGGTTGGGGACCACTCCCTGCATC 51

Oy 2712 cccctccactgagagctgtcttcttgcataaataattcttacc 2761

Db 50 CCCTCTCAGTGAAGAGCTGTCTTCTTGTCTCAATAAATCTTTTACCC 1

RESULT 48
 AM139497/c
 LOCUS 383 bp mRNA linear EST 30-OCT-1999
 DEFINITION UI-H-BII-aad-e-08-0-UI.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone
 IMAGE:2716526 3', mRNA sequence.
 ACCESSION AM139497
 VERSION AM139497.1 GI:6144215
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 383)
 NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. cDNA library Preparation: M.B. Soares Lab Clone distribution:
 NCI-CGAP clone distribution information can be found through the
 I.M.A.G.E. Consortium/ILNI, at:
 www.bio.illn.gov/bdrg/image/image.html The following repetitive
 elements were found in this cDNA sequence: 52-178, >MER524Unknwn
 Seq primer: M13 Forward
 POLYA-ties.

FEATURES
 source
 location/Qualifiers
 1..383
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2716526"
 /clone_lib="NCI_CGAP_Sub3"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; The
 NCI_CGAP_Sub3 library is a subtracted library derived from
 the NCI_CGAP_Sub1 library, which is a subtracted library
 derived from B1. B1 constitutes a mixture of 21
 normalized or subtracted NCI_CGAP libraries: NCI_CGAP_C04
 , NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_C010,
 NCI_CGAP_C016, NCI_CGAP_Kid5, NCI_CGAP_Kid12,
 NCI_CGAP_Kid3, NCI_CGAP_Kid11, NCI_CGAP_Lym2,
 NCI_CGAP_Pr2, NCI_CGAP_C08, NCI_CGAP_C111, NCI_CGAP_Lym2,
 NCI_CGAP_Pr23, NCI_CGAP_C08, NCI_CGAP_C111, NCI_CGAP_Lym2,
 NCI_CGAP_Pr23, NCI_CGAP_Ly5, NCI_CGAP_Ly4,
 NCI_CGAP_Ly19, NCI_CGAP_G04, NCI_CGAP_G06,
 NCI_CGAP_Brn25. These 21 libraries were pooled and a
 single-stranded DNA preparation of the hybridizing mixture
 was used as a tracer in a subtractive hybridization with
 a driver whose composition is detailed below:
 NCI_CGAP_Kid3 pool 1 LAM 3334-3337, 3682-3683,
 3798-3803 (IMAGE Clonoids 1322376-1323911, 1456008-1456775
 , 1500552-1502855); NCI_CGAP_Kid5 pool 1 LAM 3338-3342
 , 3722-3725, 3776-3778 (IMAGE Clonoids 1323912-1325831,
 1471368-1472903, 1492104-1493255); NCI_CGAP_Ly5 pool 1
 LAM 3575-3582, 3851-3854 (IMAGE Clonoids 141920-1417991,
 1520904-1522439); NCI_CGAP_G04 pool 1 LAM 3164-3167,
 3716-3720, 3733-3735 (IMAGE Clonoids 1257096-1258631,
 1469064-1470983, 1475592-1476743); NCI_CGAP_Pr22 pool 1
 LAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clonoids
 985608-986758, 1101192-1101959, 121928-1220615);
 NCI_CGAP_C010 pool 1 LAM 2644-2653, 2871-2872 (IMAGE
 Clonoids 1057416-1061255, 1144584-1145351). Subtraction
 was performed as previously described (Bonaldo, Lennon &

Query Match 11.0%; Score 305; DB 9; Length 383;
 Approaches To Facilitate Gene Discovery. Genome Research
 6: 791-806.
 Best Local Similarity 99.7%; Pred. No. 3.2e-149;
 Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

BASE COUNT 107 a 69 c 107 g 100 t
 ORIGIN

Y 2409 gctgaaccccaaccttaagaagcagctccgcgctaaatcctcactgaattaga 2468
 Db 373 GGTGAAMCCCACTTAAGACTGAAGACAGTCCCGGTAATCTCTACTGAATTGAGA 314
 Y 2469 acctgtctccatttggtgtcttctcccgattgatcccaacctcactatttac 2528
 Db 313 ACCGTCTTCCCAATTGGTGTGCTTTCCTCCGATTTGATCCCAACCTTACATTATTAC 254
 Y 2529 gtatactgccttctcctaattggttttacaactgtctgtccaccttttgagtgtgc 2588
 Db 253 GTATACCTGCCCTTCTCTTAATTGGTTTTCACGTGCTGCCACCTTTGAGTGTGCC 194
 Y 2589 ttgcactactaacaatcagtcagcgtgtatctccattctgaagccataaagacc 2648
 Db 193 TTGCACTACTTACAAATCAATCAAGTATCCCTATTCTGAGCCATTAAGACCA 134
 Y 2649 gactcagctgcaagtgagagaagaataaccctctgtgagagtggtggaacctccgc 2708
 Db 133 GACTCAGCTCAGTGAAGAGAGAAATCACCCTGCTGTGGGGGTTGGGACCACTCCCTGC 74
 Y 2709 atccctctccactgagagctgtcttcttgcataaataattcttaccatc 2764
 Db 73 ATCCCTCTCAGTGAAGAGCTGTCTTCTTGTCTCAATTAATCTTTTCTACCATC 18

RESULT 49
 AM579734
 LOCUS 680 bp mRNA linear EST 16-MAR-2000
 DEFINITION RCI-HT0256-080100-012-a08 HT0256 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AM579734
 VERSION AM579734.1 GI:7254783
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 680)
 HCGP http://www.judwig.org.br/ORESTES.
 TITLE The FAPESP/LICR Human Cancer Genome Project
 JOURNAL Unpublished (1999)
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@judwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.judwig.org.br/scripts/gethtml2.pl?l=RC1&t2=RC1-HT0256-
 080100-012-a08&t3=2000-01-08&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 22
 High quality sequence stop: 661.

FEATURES
 source
 location/Qualifiers
 1..680
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

/clone.lib="Hf0256"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site_1: Sma1;
Site_2: Sma1: A mini-library was made by cloning products
derived from ORESFES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 205 a 140 c 159 g 176 t
ORIGIN

Query Match 10.5%; Score 291; DB 9; Length 680;
Best Local Similarity 99.7%; Pred. No. 8.7e-142;
Matches 341; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1990 accagagagacattcagagagagagccttatttgcagaagtgctgagcggttca 2049
|||||
Db 291 ACCAGAGACACATTCAGACGAGACAGCCTTATTTGCAAAAGTGTGACGGGCTTTA 350
QY 2050 gtgcgaagtcacacattatcagacatcagagagacacattagagataaacttatgt 2109
|||||
Db 351 GTGCGAAGTCCAACTTATCAGACATCAGAGACACCTCAGATGAAACTTTATGTGT 410
QY 2110 atagggaaatgtgttacagcctttagcagagagtcatactcagacacacagagagac 2169
|||||
Db 411 ATAGGGAAATGTGTACAGCCTTATAGCAGAGATCATACCTCATCAGACACAGAGAC 470
QY 2170 acacagatgtctgtgttcttctcagacattgtctagatacacaagtgtgagacattgtgt 2229
|||||
Db 471 ACCCACTGCTGTGCTTTTACACCACTGTGTATCAAAAGTGAGACATCTGTGTGT 530
QY 2230 gattatgcatgagacgtctgttgaagactgtatctccaccactgaaaggagaaatttc 2289
|||||
Db 531 GATTATGACATGACATGCTGTGTAAGACTGTATCTCCATCCACCTGAAGAGAAATGC 590
QY 2290 tggctcatttcagagagccctgccttcctcactgtgagatg 2331
|||||
Db 591 TGGCTCATTTTTCAGGAGCCCTGCCCTTCCCTCACTGTGGATGG 632

RESULT 50
BE798936 1041 bp mRNA linear EST 20-SEP-2000
LOCUS 601583719F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3937924 5',
DEFINITION mRNA sequence.
ACCESSION BE798936
VERSION BE798936.1 GI:10220134
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE 1 (bases 1 to 1041)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM784 row: f column: 05
High quality sequence start: 58
High quality sequence stop: 705.
Location/Qualifiers
1..1041
/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone="IMAGE:3937924"
/clone.lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pORF7; Site_1: XhoI; Site_2:
ECORI; cDNA made by oligo-dr priming. Directionally
cloned into EcorI/XhoI sites using the following 5'
adapter: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 320 a 276 c 281 g 164 t
ORIGIN

Query Match 10.5%; Score 290; DB 10; Length 1041;
Best Local Similarity 100.0%; Pred. No. 3.2e-141;
Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1364 ggaagtgtgagcaaggtttagccagaagtcacacccctcagacacttaagagacacac 1423
|||||
Db 113 GGAGTGTGAGCAAGGCTTTAGCCAGAGTCACACTCATCAGACACTTAAAGACACAC 172
QY 1424 aggaagagagccttatgtatgcagaaatgtggtgtcactttagctggaatcaaacct 1483
|||||
Db 173 AGGAGAGAAAGCCTTATGTATGACACAGATGTGGCGCTCACTTATAGCTGAATCAACCT 232
QY 1484 caaaacacacacagagacacactcaagggttaaaccttatgtctgctgtgagtgagaa 1543
|||||
Db 233 CAAACACACACAGAGACACACTCAGGGGGTTAAACCTTATGTCTGCTGGAGTGGCGGCA 292
QY 1544 gtgccttagccttaaatcaaaccttaacaaacccagaggtcagacagggaggaagacc 1603
|||||
Db 293 GTGCTTTAGGCTTAAGTCAAACTTTAAACACACAGAGGTCAACACGGGGGAGAAAGCC 352
QY 1604 attgtatgtacgagatgtgtgagcagccttaccgcggaatcaaacctga 1653
|||||
Db 353 ATTGTATGTAGCGAGTGTGGCGGAGGCTTTACCGGAATCAACCTGA 402

Search completed: May 16, 2002, 09:13:25
Job time: 3170 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2002, 08:20:45 ; Search time 4965.75 seconds
(without alignments)
11681.691 Million cell updates/sec

Title: US-09-898-556A-3
2772
Perfect score: 1 cagcgcgcttaagctggtg.....ttctaccaccctcaccct 2772
Sequence:

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size: 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: listing first 1000 summaries

Database:

GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	ID	Description
--------	-------	-------	-------	--------	----	-------------

1	2772	100.0	2772	9	AB013897	AB013897 Homo sapi
2	2383	86.0	2854	9	AK056243	AK056243 Homo sapi
3	2244	81.0	135173	9	AC008806	AC008806 Homo sapi
4	2244	81.0	253217	2	AC016590	AC016590 Homo sapi
5	2193	79.1	38173	9	AC093235	AC093235 Homo sapi
6	1771	63.9	1998	9	BC004513	BC004513 Homo sapi
7	672	24.2	672	6	HUMHMR1	M20675 Human Krupp
8	571	20.6	678	6	AX067361	AX067361 Sequence
9	37	1.3	130351	2	AC024911	AC024911 Homo sapi
10	37	1.3	161625	2	AC096898	AC096898 Homo sapi
11	37	1.3	164264	2	AC034197	AC034197 Homo sapi
12	37	1.3	165379	2	AC069265	AC069265 Homo sapi
13	37	1.3	184375	2	AC026219	AC026219 Homo sapi
14	36	1.3	148270	2	AC027579	AC027579 Homo sapi
15	36	1.3	174405	2	AC022727	AC022727 Homo sapi
16	36	1.3	179792	2	AP001531	AP001531 Homo sapi
17	34	1.2	150794	2	AC025211	AC025211 Homo sapi
18	34	1.2	166650	2	AC087493	AC087493 Homo sapi
19	34	1.2	218922	2	AC008106	AC008106 Homo sapi
20	32	1.2	162453	2	AC027607	AC027607 Homo sapi
21	32	1.2	171279	2	AC036185	AC036185 Homo sapi
22	32	1.2	171601	2	AC026269	AC026269 Homo sapi
23	32	1.2	173133	2	AC016166	AC016166 Homo sapi
24	32	1.2	174701	2	AC098581	AC098581 Homo sapi
25	32	1.2	175833	2	AC073412	AC073412 Homo sapi
26	32	1.2	207362	2	AC007380	AC007380 Homo sapi
27	31	1.1	154746	9	AL157944	AL157944 Human DNA
28	31	1.1	168736	2	AC019042	AC019042 Homo sapi
29	31	1.1	177864	9	AC024996	AC024996 Homo sapi
30	29	1.0	2328	9	HSW800283	AL049964 Homo sapi
31	29	1.0	2790	9	AK056466	AK056466 Homo sapi
32	29	1.0	3178	9	AK056184	AK056184 Homo sapi
33	29	1.0	3273	9	BC021298	BC021298 Homo sapi
34	29	1.0	127867	9	HS694814	AL031673 Human DNA
35	29	1.0	170129	2	AC013770	AC013770 Homo sapi
36	28	1.0	181770	2	AC0109473	AC0109473 Homo sapi
37	28	1.0	2823	9	BC011862	BC011862 Homo sapi
38	28	1.0	3055	9	AK055734	AK055734 Homo sapi
39	28	1.0	61452	2	AC091936	AC091936 Homo sapi
40	28	1.0	62002	2	AF161800	AF161800 Homo sapi
41	28	1.0	65236	2	AC105030	AC105030 Homo sapi
42	28	1.0	80846	9	AL512489	AL512489 Human DNA
43	28	1.0	90220	9	HSJ734P14	AL049650 Human DNA
44	28	1.0	109217	9	AF159227	AF159227 Homo sapi
45	28	1.0	131328	9	AC099798	AC099798 Homo sapi
46	28	1.0	143799	9	AC011059	AC011059 Homo sapi
47	28	1.0	146831	9	AC010998	AC010998 Homo sapi
48	28	1.0	152393	2	AC021765	AC021765 Homo sapi
49	28	1.0	154504	2	AC104974	AC104974 Homo sapi
50	28	1.0	154987	9	AC080121	AC080121 Homo sapi
51	28	1.0	157402	9	AC092333	AC092333 Homo sapi
52	28	1.0	157478	2	AC015515	AC015515 Homo sapi
53	28	1.0	158039	2	AC021840	AC021840 Homo sapi
54	28	1.0	158113	2	AC011159	AC011159 Homo sapi
55	28	1.0	158999	2	AC083843	AC083843 Homo sapi
56	28	1.0	159698	9	AC027269	AC027269 Homo sapi
57	28	1.0	159956	9	AC012642	AC012642 Homo sapi
58	28	1.0	161274	9	AC012690	AC012690 Homo sapi
59	28	1.0	162671	2	AC034299	AC034299 Homo sapi
60	28	1.0	164179	9	AC012060	AC012060 Homo sapi
61	28	1.0	167073	9	AC018465	AC018465 Homo sapi
62	28	1.0	191370	30	AC017018	AC017018 Homo sapi
63	28	1.0	194236	2	AL645944	AL645944 Homo sapi
64	28	1.0	203563	2	AC069454	AC069454 Homo sapi
65	28	1.0	287571	2	AC092338	AC092338 Homo sapi
66	28	1.0	297201	2	AC091956	AC091956 Homo sapi
67	27	1.0	1667	9	HSZNF169S2	U28251 Human Krupp
68	27	1.0	1945	9	BC008213	BC008213 Homo sapi
69	27	1.0	3099	9	AF114817	AF114817 Homo sapi
70	27	1.0	3373	9	BC006247	BC006247 Homo sapi
71	27	1.0	3377	9	BC005859	BC005859 Homo sapi
72	27	1.0	25560	9	AP000298	AP000298 Homo sapi
73	27	1.0	40791	9	AF039907	AF039907 Homo sapi

74	27	1.0	100000	9	AP000044	AP000044 Homo sapi	147	25	0.9	182168	2	AL133542	AL133542 Homo sapi
75	27	1.0	100000	9	AP000112	AP000112 Homo sapi	148	25	0.9	186839	2	CNS01DXD	AL1339227 Homo sapi
76	27	1.0	100000	9	AP000188	AP000188 Homo sapi	c 149	25	0.9	190005	9	AC068061	AC068061 Homo sapi
77	27	1.0	105779	9	AL137007	AL137007 Human DNA	150	25	0.9	194055	2	AC109480	AC109480 Homo sapi
78	27	1.0	137848	9	AC097479	AC097479 Homo sapi	c 151	25	0.9	197992	2	AC020715	AC020715 Homo sapi
79	27	1.0	141843	2	AC105267	AC105267 Homo sapi	152	25	0.9	210107	9	AL365214	AL365214 Homo sapi
80	27	1.0	145294	9	AC009464	AC009464 Homo sapi	153	25	0.9	340000	9	AP001693	AP001693 Homo sapi
81	27	1.0	151458	30	AC021217	AC021217 Homo sapi	c 154	25	0.9	340000	9	AP001715	AP001715 Homo sapi
82	27	1.0	153619	9	AC026239	AC026239 Homo sapi	c 155	24	0.9	2021	5	AL053297	AL053297 Homo sapi
83	27	1.0	153851	2	AC107419	AC107419 Homo sapi	156	24	0.9	2284	9	AL057453	AL057453 Xenopus lae
84	27	1.0	155976	2	AC055113	AC055113 Homo sapi	157	24	0.9	2903	9	AE275816	AE275816 Homo sapi
85	27	1.0	172177	2	AC104190	AC104190 Homo sapi	c 158	24	0.9	26344	9	AP000202	AP000202 Homo sapi
86	27	1.0	183928	2	AL160055	AL160055 Homo sapi	159	24	0.9	58384	9	AP000241	AP000241 Homo sapi
87	27	1.0	184664	9	AC096757	AC096757 Homo sapi	160	24	0.9	67870	2	AC083787	AC083787 Homo sapi
88	27	1.0	222794	9	AC009365	AC009365 Homo sapi	161	24	0.9	71067	2	AE270510	AE270510 Homo sapi
89	27	1.0	340000	9	AP001716	AP001716 Homo sapi	c 162	24	0.9	83412	2	AC106778	AC106778 Homo sapi
90	26	0.9	50343	2	AC091190	AC091190 Homo sapi	c 163	24	0.9	85151	9	AL596219	AL596219 Homo sapi
91	26	0.9	66748	2	AC090817	AC090817 Homo sapi	164	24	0.9	85766	9	AL391559	AL391559 Human DNA
92	26	0.9	11927	2	AC005062	AC005062 Homo sapi	c 165	24	0.9	86156	2	AL162262	AL162262 Homo sapi
93	26	0.9	145495	2	CNS01D23	CNS01D23 Homo sapi	c 166	24	0.9	86371	9	AL355672	AL355672 Human DNA
94	26	0.9	153624	2	CNS01D54	CNS01D54 Homo sapi	167	24	0.9	99257	9	AL355230	AL355230 Human DNA
95	26	0.9	154199	2	AL512885	AL512885 Human DNA	c 168	24	0.9	101618	9	AC068863	AC068863 Homo sapi
96	26	0.9	165590	2	AL512885	AL512885 Human DNA	c 169	24	0.9	107033	2	AC096949	AC096949 Homo sapi
97	26	0.9	166192	9	AC096737	AC096737 Homo sapi	170	24	0.9	107898	9	AF124731	AF124731 Homo sapi
98	26	0.9	169402	2	AC027715	AC027715 Homo sapi	c 171	24	0.9	126344	9	AP000097	AP000097 Homo sapi
99	26	0.9	170378	2	AC013267	AC013267 Homo sapi	c 172	24	0.9	126590	2	AP001112	AP001112 Homo sapi
100	26	0.9	171535	9	CNS01D32	CNS01D32 Homo sapi	c 173	24	0.9	136371	9	AC025451	AC025451 Homo sapi
101	26	0.9	176872	9	AC022302	AC022302 Homo sapi	c 174	24	0.9	138224	9	AC096562	AC096562 Homo sapi
102	26	0.9	178348	9	AP000501	AP000501 Homo sapi	c 175	24	0.9	139930	2	AL160009	AL160009 Homo sapi
103	26	0.9	185209	2	AL390314	AL390314 Homo sapi	c 176	24	0.9	141488	9	AC010256	AC010256 Homo sapi
104	26	0.9	186271	2	AC024168	AC024168 Homo sapi	c 177	24	0.9	146469	2	AC068330	AC068330 Homo sapi
105	26	0.9	186735	2	AC093207	AC093207 Homo sapi	c 178	24	0.9	148697	9	AC010432	AC010432 Homo sapi
106	26	0.9	188921	2	AC025122	AC025122 Homo sapi	179	24	0.9	160042	2	AC027757	AC027757 Homo sapi
107	26	0.9	198084	2	AC019339	AC019339 Homo sapi	c 180	24	0.9	160274	2	AC073411	AC073411 Homo sapi
108	26	0.9	200426	9	AC093117	AC093117 Homo sapi	181	24	0.9	160457	9	AC006016	AC006016 Homo sapi
109	26	0.9	228434	2	AC012205	AC012205 Homo sapi	182	24	0.9	161222	2	AC016037	AC016037 Homo sapi
110	26	0.9	243887	2	AC024898	AC024898 Homo sapi	183	24	0.9	162805	2	AC020565	AC020565 Homo sapi
111	25	0.9	51	6	AX165270	AX165270 Sequence	184	24	0.9	163891	9	AC092647	AC092647 Homo sapi
112	25	0.9	51	6	AX165674	AX165674 Sequence	c 185	24	0.9	165616	9	AC073333	AC073333 Homo sapi
113	25	0.9	3104	9	AK027183	AK027183 Homo sapi	c 186	24	0.9	166554	2	AC067963	AC067963 Homo sapi
114	25	0.9	40018	9	HS10L11	HS10L11 Homo sapi	c 187	24	0.9	169516	2	AC021379	AC021379 Homo sapi
115	25	0.9	58445	9	AL536735	AL536735 Human DNA	188	24	0.9	169586	2	AC078883	AC078883 Homo sapi
116	25	0.9	59951	9	AL536735	AL536735 Human DNA	189	24	0.9	170847	2	AF188028	AF188028 Homo sapi
117	25	0.9	66129	9	AP000281	AP000281 Homo sapi	190	24	0.9	172613	9	AC087667	AC087667 Homo sapi
118	25	0.9	89323	9	AP000021	AP000021 Homo sapi	c 191	24	0.9	172759	2	AC068975	AC068975 Homo sapi
119	25	0.9	97433	2	AL512789	AL512789 Homo sapi	c 192	24	0.9	175697	2	AP002512	AP002512 Homo sapi
120	25	0.9	100000	9	AP000039	AP000039 Homo sapi	193	24	0.9	178064	2	AC023311	AC023311 Homo sapi
121	25	0.9	100000	9	AP000085	AP000085 Homo sapi	194	24	0.9	179144	2	AL606469	AL606469 Homo sapi
122	25	0.9	100000	9	AP000107	AP000107 Homo sapi	195	24	0.9	181065	2	AC104989	AC104989 Homo sapi
123	25	0.9	100000	9	AP000137	AP000137 Homo sapi	196	24	0.9	181731	9	AC021678	AC021678 Homo sapi
124	25	0.9	100000	9	AP000183	AP000183 Homo sapi	197	24	0.9	183965	9	AC068764	AC068764 Homo sapi
125	25	0.9	105062	2	AC074049	AC074049 Homo sapi	c 198	24	0.9	185832	9	AL137800	AL137800 Human DNA
126	25	0.9	123818	9	HS105613	HS105613 Homo sapi	c 199	24	0.9	186128	2	AC093286	AC093286 Homo sapi
127	25	0.9	123925	9	AC003976	AC003976 Homo sapi	200	24	0.9	186629	2	AL512603	AL512603 Homo sapi
128	25	0.9	127587	2	AC005959	AC005959 Homo sapi	201	24	0.9	189169	2	AC093008	AC093008 Homo sapi
129	25	0.9	129098	9	HS46H23	HS46H23 Human DNA	c 202	24	0.9	192550	9	AC009779	AC009779 Homo sapi
130	25	0.9	130080	9	AP003160	AP003160 Homo sapi	203	24	0.9	192722	2	AL590103	AL590103 Homo sapi
131	25	0.9	136690	2	AC107976	AC107976 Homo sapi	204	24	0.9	200562	2	AC024087	AC024087 Homo sapi
132	25	0.9	143996	2	AC068859	AC068859 Homo sapi	c 205	24	0.9	205199	2	AC090573	AC090573 Homo sapi
133	25	0.9	145325	9	AL137789	AL137789 Human DNA	c 206	24	0.9	208699	2	AC018675	AC018675 Homo sapi
134	25	0.9	148876	9	AC011455	AC011455 Homo sapi	c 207	24	0.9	215861	9	HS45537	HS45537 Homo sapi
135	25	0.9	150829	2	AC009092	AC009092 Homo sapi	c 208	24	0.9	228045	2	AC093298	AC093298 Homo sapi
136	25	0.9	151798	9	HS435C23	HS435C23 Human DNA	c 209	24	0.9	231247	9	AC008543	AC008543 Homo sapi
137	25	0.9	158866	9	AC004674	AC004674 Homo sapi	210	24	0.9	340000	9	HS21C049	HS21C049 Homo sapi
138	25	0.9	160222	9	AP002962	AP002962 Homo sapi	211	23	0.8	261	11	G01781	G01781 human STS
139	25	0.9	161757	2	AC106000	AC106000 Homo sapi	212	23	0.8	305	11	G34057	G34057 human STS
140	25	0.9	161938	9	AL360227	AL360227 Human DNA	213	23	0.8	2103	10	MUSZFP51	DI0630 Mus muscu
141	25	0.9	162606	2	AC016348	AC016348 Homo sapi	214	23	0.8	2141	10	BC010591	BC010591 Mus muscu
142	25	0.9	166329	2	AL139187	AL139187 Human DNA	215	23	0.8	2200	10	MMU41671	U041671 Mus muscu
143	25	0.9	169611	2	AC067990	AC067990 Homo sapi	216	23	0.8	2643	9	HSU09366	HSU09366 Human zinc
144	25	0.9	173902	2	AC073364	AC073364 Homo sapi	217	23	0.8	2718	9	BC001887	BC001887 Homo sapi
145	25	0.9	179106	2	AC017059	AC017059 Homo sapi	218	23	0.8	2973	10	MMZFPTA	X63747 M. musculus
146	25	0.9	179106	2	AC017059	AC017059 Homo sapi	c 219	23	0.8	48023	9	AL662888	AL662888 Human DNA

220	23	0.8	54594	9	AC073272	AC073272 Homo sapi	293	22	0.8	439	6	AX072687	AX072687 Sequence
221	23	0.8	59270	2	AC087617	AC087617 Homo sapi	294	22	0.8	484	6	AX321745	AX321745 Sequence
222	23	0.8	63037	2	AC090194	AC090194 Homo sapi	295	22	0.8	791	9	HSU20648	U20648 Human zinc
223	23	0.8	68589	2	AC087682	AC087682 Homo sapi	296	22	0.8	1553	10	AF332090	AF332090 Mus muscu
224	23	0.8	70204	2	AC010737	AC010737 Homo sapi	297	22	0.8	1557	10	AF332089	AF332089 Mus muscu
225	23	0.8	71822	2	AC027041	AC027041 Homo sapi	298	22	0.8	1739	9	HSNCF101	U00776 Human neutr
226	23	0.8	74690	2	AC022581	AC022581 Homo sapi	299	22	0.8	2295	10	AF226870	AF226870 Mus muscu
227	23	0.8	78603	2	AC022581	AC022581 Homo sapi	300	22	0.8	2382	9	AF14816	AF14816 Homo sapi
228	23	0.8	81417	2	HS1188A21	HS1188A21 Homo sapi	301	22	0.8	2411	10	AF277902	AF277902 Rattus no
229	23	0.8	89035	2	HS900012	HS900012 Homo sapi	302	22	0.8	2432	9	HSU09413	U09413 Human zinc
230	23	0.8	90507	9	HSJ324N14	HSJ324N14 Homo sapi	303	22	0.8	2679	6	AX274918	AX274918 Sequence
231	23	0.8	102505	9	HS43C13	HS43C13 Homo sapi	304	22	0.8	3031	10	AF167320	AF167320 Mus muscu
232	23	0.8	106233	9	AL359032-1	AL359032-1 Homo sapi	305	22	0.8	3146	9	AF154846	AF154846 Homo sapi
233	23	0.8	110000	2	AC024576	AC024576 Homo sapi	306	22	0.8	3271	9	BC007256	BC007256 Homo sapi
234	23	0.8	110271	9	AC006999	AC006999 Homo sapi	307	22	0.8	5541	9	AK000086	AK000086 Homo sapi
235	23	0.8	112878	2	AC084086	AC084086 Homo sapi	308	22	0.8	35641	6	AX332256	AX332256 Sequence
236	23	0.8	115296	2	AF214634	AF214634 Homo sapi	309	22	0.8	35641	6	HUMERGC46	U76566 Homo sapien
237	23	0.8	128444	2	AC005909	AC005909 Homo sapi	310	22	0.8	41601	9	AC004017	AC004017 Homo sapi
238	23	0.8	137557	2	AC108020	AC108020 Homo sapi	311	22	0.8	43958	9	AC021011	AC021011 Homo sapi
239	23	0.8	142544	2	AP002425	AP002425 Homo sapi	312	22	0.8	47466	8	AP002061	AP002061 Arabidops
240	23	0.8	144234	2	AP001570	AP001570 Homo sapi	313	22	0.8	63724	9	AL158081	AL158081 Human DNA
241	23	0.8	144279	2	AC015503	AC015503 Homo sapi	314	22	0.8	84114	9	AC003006	AC003006 Human DNA
242	23	0.8	145722	2	AC015988	AC015988 Homo sapi	315	22	0.8	88509	9	HSJ159G19	HSJ159G19 Human DNA
243	23	0.8	150113	30	AC027502	AC027502 Homo sapi	316	22	0.8	97097	9	AL353740	AL353740 Human DNA
244	23	0.8	155163	9	AC011091	AC011091 Homo sapi	317	22	0.8	103601	9	CNS01DUU	AL353740 Human Chr
245	23	0.8	155313	2	HSJ568F9	HSJ568F9 Homo sapi	318	22	0.8	110000	2	AC008749	AC008749 Homo sapi
246	23	0.8	155432	9	AP001205	AP001205 Homo sapi	319	22	0.8	112186	9	HSJ280L13	AL353740 Human DNA
247	23	0.8	155731	9	CNS05TCT	CNS05TCT Homo sapi	320	22	0.8	119453	2	AC004006	AC004006 Homo sapi
248	23	0.8	155827	9	AC011465	AC011465 Homo sapi	321	22	0.8	124154	2	AC093217	AC093217 Homo sapi
249	23	0.8	156997	9	AC025996	AC025996 Homo sapi	322	22	0.8	126795	2	AC010266	AC010266 Homo sapi
250	23	0.8	157564	2	AC104664	AC104664 Homo sapi	323	22	0.8	133418	9	AL353709	AL353709 Human DNA
251	23	0.8	158145	2	CNS07EFC	CNS07EFC Homo sapi	324	22	0.8	138752	9	AL161627	AL161627 Human DNA
252	23	0.8	158357	2	AC073494	AC073494 Homo sapi	325	22	0.8	140053	2	AC026745	AC026745 Homo sapi
253	23	0.8	160640	2	AC079295	AC079295 Homo sapi	326	22	0.8	141372	2	AC011779	AC011779 Homo sapi
254	23	0.8	163559	2	AL353765	AL353765 Homo sapi	327	22	0.8	142037	2	AC025171	AC025171 Homo sapi
255	23	0.8	164831	2	AC011061	AC011061 Homo sapi	328	22	0.8	151272	2	AC009763	AC009763 Homo sapi
256	23	0.8	166009	2	AL135785	AL135785 Human DNA	329	22	0.8	151791	2	AC108695	AC108695 Homo sapi
257	23	0.8	166207	9	AC016970	AC016970 Homo sapi	330	22	0.8	152050	9	AC079408	AC079408 Homo sapi
258	23	0.8	166966	9	AF217246	AF217246 Homo sapi	331	22	0.8	155837	2	AL445495	AL445495 Human DNA
259	23	0.8	168042	2	AC025931	AC025931 Homo sapi	332	22	0.8	156181	2	AC011677	AC011677 Homo sapi
260	23	0.8	168217	9	AL138920	AL138920 Human DNA	333	22	0.8	157812	2	AC055605	AC055605 Homo sapi
261	23	0.8	168822	9	AC007511	AC007511 Homo sapi	334	22	0.8	159250	9	AC073528	AC073528 Homo sapi
262	23	0.8	171179	2	AC034140	AC034140 Homo sapi	335	22	0.8	159250	9	AC104775	AC104775 Homo sapi
263	23	0.8	171545	2	AC008394	AC008394 Homo sapi	336	22	0.8	159898	2	AC063974	AC063974 Homo sapi
264	23	0.8	174080	2	AF217246	AF217246 Homo sapi	337	22	0.8	159998	2	AC011566	AC011566 Homo sapi
265	23	0.8	176213	9	AC009518	AC009518 Human Chr	338	22	0.8	163198	2	AL161634	AL161634 Homo sapi
266	23	0.8	177286	9	AC008085	AC008085 Homo sapi	339	22	0.8	164220	30	AC024130	AC024130 Homo sapi
267	23	0.8	177522	9	AC093566	AC093566 Homo sapi	340	22	0.8	164479	9	AC103922	AC103922 Homo sapi
268	23	0.8	179238	2	AC009451	AC009451 Homo sapi	341	22	0.8	165338	9	AC093805	AC093805 Homo sapi
269	23	0.8	179876	2	AC013277	AC013277 Homo sapi	342	22	0.8	165589	2	AL356584	AL356584 Human DNA
270	23	0.8	182506	9	AL359077	AL359077 Human DNA	343	22	0.8	165608	2	AC018473	AC018473 Homo sapi
271	23	0.8	183559	2	AC073501	AC073501 Homo sapi	344	22	0.8	166983	2	AC002541	AC002541 Homo sapi
272	23	0.8	184510	2	AC018894	AC018894 Homo sapi	345	22	0.8	167330	2	AC021822	AC021822 Homo sapi
273	23	0.8	185322	2	AC025667	AC025667 Mus muscu	346	22	0.8	167943	9	HS267W20	HS267W20 Homo sapi
274	23	0.8	186208	30	AC019204	AC019204 Homo sapi	347	22	0.8	168779	9	AC026405	AC026405 Homo sapi
275	23	0.8	187253	2	AC104186	AC104186 Homo sapi	348	22	0.8	169089	9	AC008751	AC008751 Homo sapi
276	23	0.8	188992	2	AC073873	AC073873 Homo sapi	349	22	0.8	169418	2	AC092046	AC092046 Homo sapi
277	23	0.8	189098	2	AP001491	AP001491 Homo sapi	350	22	0.8	169418	2	AC092046	AC092046 Homo sapi
278	23	0.8	189241	2	AC090402	AC090402 Homo sapi	351	22	0.8	169849	2	AC108715	AC108715 Homo sapi
279	23	0.8	189370	2	AC022254	AC022254 Homo sapi	352	22	0.8	171495	2	AC027571	AC027571 Homo sapi
280	23	0.8	191832	2	AC011104	AC011104 Homo sapi	353	22	0.8	172687	2	AC027144	AC027144 Homo sapi
281	23	0.8	193057	9	AC022325	AC022325 Homo sapi	354	22	0.8	172720	2	AC091886	AC091886 Homo sapi
282	23	0.8	194851	9	AC079264	AC079264 Homo sapi	355	22	0.8	173917	2	AC027476	AC027476 Homo sapi
283	23	0.8	197143	2	AC069173	AC069173 Homo sapi	356	22	0.8	174342	2	AC016901	AC016901 Homo sapi
284	23	0.8	198742	2	AL355360	AL355360 Homo sapi	357	22	0.8	175552	2	AC027475	AC027475 Homo sapi
285	23	0.8	199534	2	AC025959	AC025959 Homo sapi	358	22	0.8	180925	2	AL0603926	AL0603926 Homo sapi
286	23	0.8	200820	2	AC107896	AC107896 Homo sapi	359	22	0.8	181438	2	AC023047	AC023047 Homo sapi
287	23	0.8	201344	2	AC108496	AC108496 Homo sapi	360	22	0.8	183072	2	AC026338	AC026338 Homo sapi
288	23	0.8	205195	2	AC025963	AC025963 Homo sapi	361	22	0.8	183215	2	AC023370	AC023370 Homo sapi
289	23	0.8	207757	2	AL589784	AL589784 Homo sapi	362	22	0.8	183228	9	AC010547	AC010547 Homo sapi
290	23	0.8	209958	2	AC087135	AC087135 Mus muscu	363	22	0.8	184181	9	AL158141	AL158141 Human DNA
291	22	0.8	282	9	HMZNFEN	HMZNFEN Homo sapien	364	22	0.8	185322	2	AC025667	AC025667 Mus muscu
292	22	0.8	292	9	HSZNFPT3	HSZNFPT3 Homo sapien	365	22	0.8	185622	9	AC091667	AC091667 Homo sapi

366	22	0.8	187628	9	AC011311	Homo sapi	C 439	21	0.8	140591	2	AC090888	AC090888	Homo sapi
C 367	22	0.8	188629	9	AC084365	Homo sapi	C 440	21	0.8	141441	9	AC087073	AC087073	Homo sapi
368	22	0.8	190499	2	AC068392	Homo sapi	C 441	21	0.8	143794	9	AC009315	AC009315	Homo sapi
369	22	0.8	190907	2	AC005261	Homo sapi	C 442	21	0.8	144374	2	AC009283	AC009283	Homo sapi
C 370	22	0.8	191960	9	CNS01R68	Homo sapi	C 443	21	0.8	145399	2	AF228728	AF228728	Homo sapi
C 371	22	0.8	192529	2	AC019190	Homo sapi	C 444	21	0.8	145408	2	AC024303	AC024303	Homo sapi
C 372	22	0.8	193041	9	AC092017	Homo sapi	C 445	21	0.8	147031	2	AC084264	AC084264	Homo sapi
C 373	22	0.8	195358	2	AC040892	Homo sapi	C 446	21	0.8	151202	9	AL160168	AL160168	Homo sapi
C 374	22	0.8	195653	9	AC005873	ctdb_15_o	C 447	21	0.8	153704	2	AC008567	AC008567	Homo sapi
C 375	22	0.8	196319	9	CNS01DP4	Human chr	C 448	21	0.8	154600	2	AC110610	AC110610	Homo sapi
C 376	22	0.8	198751	9	AC022148	Homo sapi	C 449	21	0.8	154917	9	AC026185	AC026185	Homo sapi
C 377	22	0.8	202077	2	AC073911	Homo sapi	C 450	21	0.8	155459	2	AC016619	AC016619	Homo sapi
C 378	22	0.8	204924	2	AC015616	Homo sapi	C 451	21	0.8	155359	2	AC026477	AC026477	Homo sapi
C 379	22	0.8	205815	2	AC008759	Homo sapi	C 452	21	0.8	156045	2	AC026710	AC026710	Homo sapi
C 380	22	0.8	206346	2	AC027438	Homo sapi	C 453	21	0.8	156279	2	AC027707	AC027707	Homo sapi
C 381	22	0.8	209958	2	AC087135	Mus muscu	C 454	21	0.8	156681	2	AC104818	AC104818	Homo sapi
C 382	22	0.8	211259	2	AC024928	Homo sapi	C 455	21	0.8	157300	10	AL592522	AL592522	Mouse DNA
C 383	22	0.8	218233	2	AC099669	Homo sapi	C 456	21	0.8	157349	2	AF145206	AF145206	Homo sapi
C 384	22	0.8	230864	2	AC079523	Mus muscu	C 457	21	0.8	158213	2	AF214635	AF214635	Homo sapi
C 385	22	0.8	231146	2	AL669941	Mus muscu	C 458	21	0.8	158724	9	AL160010	AL160010	Human DNA
C 386	22	0.8	236029	2	AC019232	Homo sapi	C 459	21	0.8	160182	2	AC026594	AC026594	Homo sapi
C 387	22	0.8	245235	2	AL611964	Homo sapi	C 460	21	0.8	160334	2	AC079549	AC079549	Homo sapi
C 388	21	0.8	392	9	HUMNR70H10	Homo sapi	C 461	21	0.8	160615	9	AC026786	AC026786	Homo sapi
C 389	21	0.8	444	9	HSZNF19S1	Human chr	C 462	21	0.8	161078	9	AP003350	AP003350	Homo sapi
C 390	21	0.8	538	5	EMZNF6	Human chr	C 463	21	0.8	161286	2	AC025120	AC025120	Homo sapi
C 391	21	0.8	538	5	EMZNF6M	Human chr	C 464	21	0.8	161649	2	AC023652	AC023652	Homo sapi
C 392	21	0.8	1333	9	BC014381	Homo sapi	C 465	21	0.8	162700	2	AC024105	AC024105	Homo sapi
C 393	21	0.8	1414	9	HSK0X1	Human chr	C 466	21	0.8	162943	9	CNS01DXY	CNS01DXY	Human chr
C 394	21	0.8	1493	9	AF193057	Homo sapi	C 467	21	0.8	164374	9	AL355612	AL355612	Human DNA
C 395	21	0.8	1494	9	AF217988	Homo sapi	C 468	21	0.8	164719	2	AC068623	AC068623	Homo sapi
C 396	21	0.8	1529	9	BC007393	Homo sapi	C 469	21	0.8	165079	2	AC024517	AC024517	Homo sapi
C 397	21	0.8	1625	9	AK000223	Homo sapi	C 470	21	0.8	165693	2	AC087879	AC087879	Homo sapi
C 398	21	0.8	1663	9	AF155100	Homo sapi	C 471	21	0.8	165905	2	AC026752	AC026752	Homo sapi
C 399	21	0.8	1675	9	HSMB00838	Homo sapi	C 472	21	0.8	166862	2	AC069533	AC069533	Homo sapi
C 400	21	0.8	1960	9	HSMB01833	Homo sapi	C 473	21	0.8	166937	2	AC012255	AC012255	Homo sapi
C 401	21	0.8	2083	9	AF166122	Homo sapi	C 474	21	0.8	167275	2	AC074123	AC074123	Homo sapi
C 402	21	0.8	2203	9	AK056313	Homo sapi	C 475	21	0.8	167300	2	AC021394	AC021394	Homo sapi
C 403	21	0.8	2330	9	BC001791	Homo sapi	C 476	21	0.8	167594	2	AC095035	AC095035	Homo sapi
C 404	21	0.8	2382	6	AR170700	Sequence	C 477	21	0.8	167806	2	AC092572	AC092572	Homo sapi
C 405	21	0.8	2566	9	AK056489	Homo sapi	C 478	21	0.8	168440	2	AC105282	AC105282	Homo sapi
C 406	21	0.8	2732	9	AK027700	Homo sapi	C 479	21	0.8	168474	3	AC093454	AC093454	Drosophill
C 407	21	0.8	2864	9	AK056548	Homo sapi	C 480	21	0.8	169117	9	AC087894	AC087894	Homo sapi
C 408	21	0.8	3007	10	BC021326	Mus muscu	C 481	21	0.8	169533	2	AC108119	AC108119	Homo sapi
C 409	21	0.8	3059	9	BC006436	Homo sapi	C 482	21	0.8	169740	2	AF186193	AF186193	Homo sapi
C 410	21	0.8	3204	9	AK027363	Homo sapi	C 483	21	0.8	169944	2	AC091643	AC091643	Homo sapi
C 411	21	0.8	11000	8	AF286114	Xenothec	C 484	21	0.8	170445	2	AC068594	AC068594	Homo sapi
C 412	21	0.8	12404	9	HUMPCBD	Human DNA	C 485	21	0.8	170656	2	AP001263	AP001263	Homo sapi
C 413	21	0.8	42967	9	AC093619	Homo sapi	C 486	21	0.8	171012	2	AC073336	AC073336	Homo sapi
C 414	21	0.8	44059	9	AL669815	Human DNA	C 487	21	0.8	171548	2	AC024732	AC024732	Homo sapi
C 415	21	0.8	52346	2	AC084346	Homo sapi	C 488	21	0.8	172421	2	AC069538	AC069538	Homo sapi
C 416	21	0.8	56745	2	AC084347	Homo sapi	C 489	21	0.8	173118	2	AC078827	AC078827	Homo sapi
C 417	21	0.8	59228	2	AC069452	Homo sapi	C 490	21	0.8	173746	2	AC024040	AC024040	Homo sapi
C 418	21	0.8	74709	2	AC105479	Rattus no	C 491	21	0.8	174612	9	AC018494	AC018494	Homo sapi
C 419	21	0.8	79461	2	AC016422	Homo sapi	C 492	21	0.8	174831	2	AC008966	AC008966	Homo sapi
C 420	21	0.8	84873	2	AC024570	Homo sapi	C 493	21	0.8	175256	2	AC090775	AC090775	Homo sapi
C 421	21	0.8	95738	2	AC020272	Drosophill	C 494	21	0.8	175339	9	AC005772	AC005772	Homo sapi
C 422	21	0.8	97959	9	AC093085	Homo sapi	C 495	21	0.8	175454	2	AC073176	AC073176	Homo sapi
C 423	21	0.8	98889	2	AL590863	Homo sapi	C 496	21	0.8	175529	2	AL591916	AL591916	Homo sapi
C 424	21	0.8	103054	2	AC014803	Drosophill	C 497	21	0.8	176018	9	AC093878	AC093878	Homo sapi
C 425	21	0.8	108892	2	AC034232	Homo sapi	C 498	21	0.8	176197	2	AC092043	AC092043	Homo sapi
C 426	21	0.8	111837	9	AC069383	Homo sapi	C 499	21	0.8	177267	9	AC026116	AC026116	Homo sapi
C 427	21	0.8	112359	9	HSU60822	Human dyscr	C 500	21	0.8	177636	2	AC094891	AC094891	Homo sapi
C 428	21	0.8	113345	9	AF109076	Homo sapi	C 501	21	0.8	177834	9	AC018811	AC018811	Homo sapi
C 429	21	0.8	113654	9	AL606526	Human DNA	C 502	21	0.8	178141	2	AC074345	AC074345	Homo sapi
C 430	21	0.8	123530	9	AL139110	Human DNA	C 503	21	0.8	178184	9	AC008770	AC008770	Homo sapi
C 431	21	0.8	128386	2	AL356782	Human DNA	C 504	21	0.8	178246	2	AP003782	AP003782	Homo sapi
C 432	21	0.8	133950	2	AC008899	Homo sapi	C 505	21	0.8	178367	9	AC006615	AC006615	Homo sapi
C 433	21	0.8	134249	9	AC019116	Homo sapi	C 506	21	0.8	178377	9	AC092505	AC092505	Homo sapi
C 434	21	0.8	135345	9	AL356294	Human DNA	C 507	21	0.8	179319	2	AC109493	AC109493	Homo sapi
C 435	21	0.8	138076	2	AC099169	Rattus no	C 508	21	0.8	179581	2	AC109397	AC109397	Homo sapi
C 436	21	0.8	139476	9	AC008773	Homo sapi	C 509	21	0.8	180510	2	AC022415	AC022415	Homo sapi
C 437	21	0.8	139623	2	AC040932	Homo sapi	C 510	21	0.8	180615	9	AC025549	AC025549	Homo sapi
C 438	21	0.8	140156	2	AC027261	Homo sapi	C 511	21	0.8	180672	9	AC011451	AC011451	Homo sapi

512	21	0.8	180717	2	AC068169	AC068169 Homo sapi	585	20	0.7	2627	9	BC013741	BC013741 Homo sapi
513	21	0.8	181141	2	AC093621	Homo sapi	586	20	0.7	2640	10	MMMRVZFP	X74855 M.musculus
514	21	0.8	181312	2	AC091117	Homo sapi	587	20	0.7	2779	9	HSMB00634	AL096732 Rattus norv
515	21	0.8	181720	3	AC010114	Drosophili	588	20	0.7	2821	10	RNU56732	A1745587 Rattus norv
516	21	0.8	181858	2	AC084856	Homo sapi	589	20	0.7	2871	9	HSR245587	X84996 X.laeyvis mr
517	21	0.8	184043	2	AC092835	Homo sapi	590	20	0.7	2973	5	X1STAR	X57715 H.sapiens mr
518	21	0.8	185091	2	AC019835	Homo sapi	591	20	0.7	3759	9	HS2NF742	M8502 Mus musculus
519	21	0.8	185916	2	AC021646	Homo sapi	592	20	0.7	4582	10	MUSPLM24X	X82126 H.sapiens H
520	21	0.8	186094	2	AC084225	Homo sapi	593	20	0.7	9367	9	HSBOK2EX2	AC107199 Rattus no
521	21	0.8	186091	2	AC011242	Homo sapi	594	20	0.7	25103	2	AC094656	AC094656 Rattus no
522	21	0.8	186265	2	AC078959	Homo sapi	595	20	0.7	31938	2	AC094656	AC020951 Homo sapi
523	21	0.8	186793	3	AC010660	Drosophili	596	20	0.7	35715	9	AC004397	AC004397 Homo sapi
524	21	0.8	186868	2	AC023639	Homo sapi	597	20	0.7	36510	9	AC003961	AC003961 Human Cos
525	21	0.8	187415	2	AC026352	Homo sapi	598	20	0.7	39526	9	CEFE52B11	Z82268 Caenorhabdi
526	21	0.8	187746	2	AC087123	Mus muscu	599	20	0.7	42861	3	AC105042	AC105042 Mus muscu
527	21	0.8	188734	2	AC069443	Homo sapi	600	20	0.7	50396	2	AC105042	AC105042 Homo sapi
528	21	0.8	190268	2	AC108222	Homo sapi	601	20	0.7	51099	2	AC099646_3	Continuation (4 of
529	21	0.8	192963	9	AL354763	Human DNA	602	20	0.7	55823	2	AC040923	AC040923 Homo sapi
530	21	0.8	193935	2	AC020609	Homo sapi	603	20	0.7	57145	9	AL391820	AL391820 Human DNA
531	21	0.8	195280	2	AC025081	Homo sapi	604	20	0.7	57287	9	AC008283	AC008283 Homo sapi
532	21	0.8	197425	2	AC016776	Homo sapi	605	20	0.7	57594	2	AC104552	AC104552 Mus muscu
533	21	0.8	198586	2	AC092503	Homo sapi	606	20	0.7	59712	2	AC104221	AC104221 Mus muscu
534	21	0.8	201437	2	AC015981	Homo sapi	607	20	0.7	62716	2	AC012201	AC012201 Homo sapi
535	21	0.8	203006	9	HS1145L23	Human DNA	608	20	0.7	64008	2	AC105101	AC105101 Homo sapi
536	21	0.8	207908	2	AC108668	Homo sapi	609	20	0.7	64365	2	AC099842	AC099842 Homo sapi
537	21	0.8	218233	2	AC099669	Homo sapi	610	20	0.7	66975	9	AL135791	AL135791 Human DNA
538	21	0.8	218445	2	AC083959	Homo sapi	611	20	0.7	73198	9	HSB127L3	HSB127L3 Homo sapi
539	21	0.8	218670	9	AC096669	Homo sapi	612	20	0.7	74105	2	AC087648	AC087648 Homo sapi
540	21	0.8	218977	9	AC007500	Homo sapi	613	20	0.7	74645	2	AC009592	AC009592 Homo sapi
541	21	0.8	219658	9	AC022681	Homo sapi	614	20	0.7	78667	2	AC023522	AC023522 Homo sapi
542	21	0.8	223791	2	AC084042	Mus muscu	615	20	0.7	86719	9	HSJ906P16	HSJ906P16 Human DNA
543	21	0.8	223805	2	AC079475	Mus muscu	616	20	0.7	87903	9	HS137B17	HS137B17 Human DNA
544	21	0.8	226116	2	AC084078	Homo sapi	617	20	0.7	93643	9	HS629A15	HS629A15 Human DNA
545	21	0.8	226860	2	AL359920	Human DNA	618	20	0.7	103367	9	HSB2A2P16	HSB2A2P16 Human DNA
546	21	0.8	229687	2	AC107068	Homo sapi	619	20	0.7	105412	2	AC011458	AC011458 Homo sapi
547	21	0.8	241714	2	AC083834	Mus muscu	620	20	0.7	106341	9	AL390316	AL390316 Human DNA
548	21	0.8	248940	2	AC103563	Homo sapi	621	20	0.7	108542	9	AC104562	AC104562 Homo sapi
549	21	0.8	250681	2	AP004282	Homo sapi	622	20	0.7	108542	2	AF263550	AF263550 Homo sapi
550	21	0.8	257383	2	AC079494	Mus muscu	623	20	0.7	110000	2	AL135900	AL135900 Homo sapi
551	21	0.8	274656	2	AC079538	Mus muscu	624	20	0.7	110000	2	AL596328_3	Continuation (4 of
552	21	0.8	277603	2	AC079543	Mus muscu	625	20	0.7	113022	2	AC084846	AC084846 Homo sapi
553	21	0.8	283821	3	ABE03528	Drosophili	626	20	0.7	113284	2	AC074300	AC074300 Homo sapi
554	21	0.8	344150	9	AF235103	Homo sapi	627	20	0.7	119034	9	AC007544	AC007544 Homo sapi
555	20	0.7	99	10	RNU78142	Rattus norv	628	20	0.7	120450	9	AC093393	AC093393 Homo sapi
556	20	0.7	413	6	AX072834	Sequence	629	20	0.7	120769	2	AC092120	AC092120 Homo sapi
557	20	0.7	426	11	G58442	SHGC-104386	630	20	0.7	127058	2	AC025652	AC025652 Homo sapi
558	20	0.7	429	11	G33231	human STS S	631	20	0.7	127606	2	AC110073	AC110073 Homo sapi
559	20	0.7	888	6	AX188154	Sequence	632	20	0.7	128779	9	HS29X1	Z98745 Human DNA
560	20	0.7	1094	9	HS2NF74A2	Homo sapi	633	20	0.7	130067	9	AC007228	AC007228 Homo sapi
561	20	0.7	1106	5	AF026084	Taeniolopy	634	20	0.7	131215	9	AC079614	AC079614 Homo sapi
562	20	0.7	1263	9	BC022246	Homo sapi	635	20	0.7	132242	9	AC069205	AC069205 Homo sapi
563	20	0.7	1674	9	AK054647	Homo sapi	636	20	0.7	135432	2	AC069036	AC069036 Homo sapi
564	20	0.7	1699	9	BC007868	Homo sapi	637	20	0.7	136492	2	AC011143	AC011143 Homo sapi
565	20	0.7	1776	5	AF242519	Homo sapi	638	20	0.7	136725	9	AC026369	AC026369 Homo sapi
566	20	0.7	1800	5	XLXNA	Xenopus lae	639	20	0.7	141579	2	AC044793	AC044793 Homo sapi
567	20	0.7	1892	6	AR116036	Sequence	640	20	0.7	142000	9	AC095353	AC095353 Homo sapi
568	20	0.7	2080	10	BC010442	Mus muscu	641	20	0.7	142796	2	AC006411	AC006411 Homo sapi
569	20	0.7	2181	6	AB057740	Ciona sav	642	20	0.7	146471	2	AC026545	AC026545 Homo sapi
570	20	0.7	2250	10	AX274847	Sequence	643	20	0.7	146768	9	AC016194	AC016194 Homo sapi
571	20	0.7	2250	10	BC006587	Mus muscu	644	20	0.7	147185	9	AC020589	AC020589 Homo sapi
572	20	0.7	2356	10	MOSK1ZF	Mus muscu	645	20	0.7	147706	9	AC092578	AC092578 Homo sapi
573	20	0.7	2382	10	MMY14295	Mus muscu	646	20	0.7	147876	9	AC024600	AC024600 Homo sapi
574	20	0.7	2386	9	HSBOK2H1	H.sapiens H	647	20	0.7	147929	9	AC016951	AC016951 Homo sapi
575	20	0.7	2407	9	AB066544	Macaca fa	648	20	0.7	147978	9	AL136370	AL136370 Human DNA
576	20	0.7	2416	9	HSZNF741	H.sapiens z	649	20	0.7	148359	9	AC007434	AC007434 Homo sapi
577	20	0.7	2416	11	G27154	SHGC-31560	650	20	0.7	148466	2	AC023167	AC023167 Mus muscu
578	20	0.7	2454	10	AF184111	Mus muscu	651	20	0.7	151777	9	AL354872	AL354872 Human DNA
579	20	0.7	2478	9	AK027592	Homo sapi	652	20	0.7	152144	2	AC016293	AC016293 Homo sapi
580	20	0.7	2504	9	HS009367	Human zinc	653	20	0.7	152144	2	AC016293	AC016293 Homo sapi
581	20	0.7	2530	10	BC006421	Homo sapi	654	20	0.7	152334	2	AL356299	AL356299 Human DNA
582	20	0.7	2554	10	BC011183	Mus muscu	655	20	0.7	152649	2	AC016353	AC016353 Homo sapi
583	20	0.7	2593	10	AB010143	Mus muscu	656	20	0.7	153241	2	AC024960	AC024960 Homo sapi
584	20	0.7	2604	10	MDZFP30	Z30174 M.domesticu	657	20	0.7	153479	2	AC095130	AC095130 Rattus no

658	20	0.7	154307	2	AC108150	AC108150 Homo sapi	731	20	0.7	186142	2	AL390124	AL390124 Homo sapi
c 659	20	0.7	155094	2	AL512637	AL512637 Homo sapi	732	20	0.7	186760	2	AC092123	AC092123 Homo sapi
660	20	0.7	156244	9	AC011815	AC011815 Homo sapi	c 733	20	0.7	187295	2	AC021531	AC021531 Homo sapi
c 661	20	0.7	156466	9	AL354796	AL354796 Human DNA	c 734	20	0.7	187543	9	AC005678	AC005678 Homo sapi
662	20	0.7	156593	2	AC078999	AC078999 Homo sapi	c 735	20	0.7	187568	9	AP002967	AP002967 Homo sapi
c 663	20	0.7	158146	2	AC108862	AC108862 Homo sapi	736	20	0.7	188330	2	AC015456	AC015456 Homo sapi
c 664	20	0.7	158152	2	AC095177	AC095177 Rattus no	737	20	0.7	188377	2	AC024894	AC024894 Homo sapi
c 665	20	0.7	158370	2	AC097167	AC097167 Rattus no	738	20	0.7	190000	2	AC027701	AC027701 Homo sapi
c 666	20	0.7	158460	2	AC098177	AC098177 Rattus no	739	20	0.7	190200	2	AC022021	AC022021 Homo sapi
c 667	20	0.7	158990	2	AL591702	AL591702 Homo sapi	c 740	20	0.7	190626	2	AC062036	AC062036 Homo sapi
668	20	0.7	159451	9	AC012618	AC012618 Homo sapi	c 741	20	0.7	190946	9	CNS01DV1	CNS01DV1 Homo sapi
c 669	20	0.7	159784	2	AC061996	AC061996 Homo sapi	742	20	0.7	192592	2	AC005500	AC005500 Homo sapi
c 670	20	0.7	159960	2	AC098620	AC098620 Rattus no	c 743	20	0.7	194052	2	AC053537	AC053537 Homo sapi
671	20	0.7	160003	2	AC021051	AC021051 Homo sapi	744	20	0.7	194297	9	AC044810	AC044810 Homo sapi
672	20	0.7	160242	2	AP002778	AP002778 Homo sapi	c 745	20	0.7	194599	2	AC073668	AC073668 Homo sapi
673	20	0.7	160836	2	AC096630	AC096630 Pan trogl	c 746	20	0.7	195156	9	AC090660	AC090660 Homo sapi
674	20	0.7	162625	2	CEN0083	Z83124 Caenorhabd	c 747	20	0.7	195446	9	AC023170	AC023170 Homo sapi
c 675	20	0.7	165486	2	AC010830	AC010830 Homo sapi	c 748	20	0.7	197328	9	AC01967	AC01967 Homo sapi
676	20	0.7	165566	2	AC092295	AC092295 Homo sapi	749	20	0.7	197784	9	AL450326	AL450326 Human DNA
677	20	0.7	166288	2	AL627095	AL627095 Homo sapi	c 750	20	0.7	197784	9	AL450326	AL450326 Human DNA
678	20	0.7	167344	2	AC096481	AC096481 Rattus no	751	20	0.7	197819	2	AC073675	AC073675 Homo sapi
c 679	20	0.7	167904	2	AL359173	AL359173 Human DNA	752	20	0.7	198127	2	AP001593	AP001593 Homo sapi
c 680	20	0.7	168870	2	AC021873	AC021873 Homo sapi	753	20	0.7	198440	2	AC009993	AC009993 Homo sapi
c 681	20	0.7	168907	2	AC018390	AC018390 Homo sapi	c 754	20	0.7	198570	9	HS343C1	HS343C1 Human DNA
c 682	20	0.7	170281	2	AC079910	AC079910 Homo sapi	755	20	0.7	198570	9	AC092378	AC092378 Homo sapi
683	20	0.7	170520	2	AL358172	AL358172 Homo sapi	756	20	0.7	200418	9	AC074138	AC074138 Homo sapi
c 684	20	0.7	170579	2	AC020985	AC020985 Homo sapi	757	20	0.7	200967	2	AC098511	AC098511 Homo sapi
c 685	20	0.7	170716	2	AC095000	AC095000 Rattus no	758	20	0.7	200991	2	AC106877	AC106877 Homo sapi
c 686	20	0.7	171428	2	AC013407	AC013407 Homo sapi	c 759	20	0.7	201371	2	AC073786	AC073786 Homo sapi
c 687	20	0.7	171779	9	AC079338	AC079338 Homo sapi	c 760	20	0.7	201495	2	AC094415	AC094415 Homo sapi
c 688	20	0.7	172350	2	AC079234	AC079234 Homo sapi	c 761	20	0.7	201895	2	AC097710	AC097710 Homo sapi
c 689	20	0.7	172533	2	AC002381	AC002381 Human BAC	c 762	20	0.7	201945	2	AL671854	AL671854 Homo sapi
c 690	20	0.7	172723	2	AL356320	AL356320 Homo sapi	c 763	20	0.7	204175	2	AC037488	AC037488 Homo sapi
691	20	0.7	172923	9	AL160255	AL160255 Human DNA	764	20	0.7	205832	2	AL390728	AL390728 Homo sapi
c 692	20	0.7	173188	2	AC023509	AC023509 Homo sapi	765	20	0.7	206660	2	AC092353	AC092353 Homo sapi
c 693	20	0.7	173701	2	AC018946	AC018946 Homo sapi	c 766	20	0.7	209456	2	AC098875	AC098875 Homo sapi
c 694	20	0.7	174429	2	AC104637	AC104637 Homo sapi	c 767	20	0.7	211463	2	AC016508	AC016508 Homo sapi
c 695	20	0.7	174533	2	AC073718	Mus muscu	c 768	20	0.7	212134	2	AC024037	AC024037 Homo sapi
c 696	20	0.7	175654	2	AC025496	AC025496 Homo sapi	c 769	20	0.7	213660	2	AC073699	AC073699 Homo sapi
c 697	20	0.7	175950	2	AC098952	AC098952 Rattus no	770	20	0.7	214180	2	AC009098	AC009098 Homo sapi
c 698	20	0.7	176512	9	AC012323	AC012323 Homo sapi	c 771	20	0.7	215584	2	AL390779	AL390779 Homo sapi
699	20	0.7	176759	2	AC037464	AC037464 Homo sapi	c 772	20	0.7	216706	9	AC090420	AC090420 Homo sapi
c 700	20	0.7	177164	2	AC103450	AC103450 Rattus no	c 773	20	0.7	218859	2	AC022912	AC022912 Homo sapi
701	20	0.7	177444	9	AC011468	AC011468 Homo sapi	c 774	20	0.7	219043	2	AC073812	AC073812 Homo sapi
702	20	0.7	178184	9	AC008770	AC008770 Homo sapi	c 775	20	0.7	220678	2	AC023857	AC023857 Homo sapi
703	20	0.7	178323	9	AL354810	AL354810 Human DNA	776	20	0.7	223192	2	AL671899	AL671899 Homo sapi
704	20	0.7	178424	9	AC069020	AC069020 Homo sapi	c 777	20	0.7	224573	2	AC008758	AC008758 Homo sapi
705	20	0.7	178624	9	AC009756	AC009756 Homo sapi	c 778	20	0.7	226178	2	AL669827	AL669827 Homo sapi
706	20	0.7	179272	2	AC069149	AC069149 Homo sapi	c 779	20	0.7	227074	2	AC060772	AC060772 Homo sapi
c 707	20	0.7	179272	2	AC069149	AC069149 Homo sapi	780	20	0.7	229000	2	AL391381	AL391381 Homo sapi
c 708	20	0.7	179380	9	AL355501	AL355501 Human DNA	c 781	20	0.7	241099	2	AC099389	AC099389 Homo sapi
709	20	0.7	179491	2	AC020673	AC020673 Homo sapi	c 782	20	0.7	242406	2	AL669845	AL669845 Homo sapi
c 710	20	0.7	179526	2	AC034133	AC034133 Homo sapi	783	20	0.7	244931	2	AC079495	AC079495 Homo sapi
711	20	0.7	179581	2	AC009397	AC009397 Homo sapi	c 784	20	0.7	267635	2	AC073772	AC073772 Homo sapi
712	20	0.7	179767	9	AC069223	AC069223 Homo sapi	785	20	0.7	272168	9	AC079561	AC079561 Homo sapi
713	20	0.7	180331	9	AC007375	AC007375 Homo sapi	786	20	0.7	275159	2	U82670	U82670 Homo sapien
714	20	0.7	180510	9	AC022415	AC022415 Homo sapi	787	20	0.7	297900	2	AC006704	AC006704 Homo sapi
c 715	20	0.7	180638	9	AC068763	AC068763 Homo sapi	788	20	0.7	300050	9	HSMX1A	HSMX1A Homo sapi
c 716	20	0.7	180727	2	AC027221	AC027221 Homo sapi	789	20	0.7	318503	2	AC074166	AC074166 Homo sapi
c 717	20	0.7	180816	2	AL627249	AL627249 Dantio ier	c 790	20	0.7	318503	2	AC074166	AC074166 Homo sapi
718	20	0.7	180900	2	AC018975	AC018975 Homo sapi	c 791	20	0.7	340000	9	HS21C084	HS21C084 Homo sapi
719	20	0.7	180964	2	AL591642	AL591642 Homo sapi	c 792	20	0.7	308	9	HS261Z9	HS261Z9 Homo sapi
720	20	0.7	181026	2	AC009855	AC009855 Homo sapi	793	19	0.7	326	11	G23841	G23841 human STS
721	20	0.7	181421	9	AC011637	AC011637 Homo sapi	794	19	0.7	357	3	AP077675	AP077675 Strongylo
c 722	20	0.7	181451	2	AP003038	AP003038 Homo sapi	795	19	0.7	375	11	G34024	G34024 human STS
723	20	0.7	181709	2	AL672031	Mus muscu	796	19	0.7	375	11	HUM4S1258	HUM4S1258 human chrom
724	20	0.7	181773	2	AL365365	AL365365 Homo sapi	c 797	19	0.7	378	9	HS1607004	HS1607004 human STS
725	20	0.7	182125	2	AC090105	AC090105 Homo sapi	798	19	0.7	387	11	G26051	G26051 human STS
726	20	0.7	182188	2	AC048353	AC048353 Homo sapi	c 799	19	0.7	397	11	G60159	G60159 SHGC-131044
727	20	0.7	182617	2	AC007731	AC007731 Homo sapi	c 800	19	0.7	412	11	HSPE53D08	HSPE53D08 human STS
c 728	20	0.7	184439	2	AP003085	AP003085 Homo sapi	c 801	19	0.7	462	11	G51998	G51998 SHGC-81978
c 729	20	0.7	185493	2	AC068986	AC068986 Homo sapi	802	19	0.7	483	11	G28077	G28077 human STS
c 730	20	0.7	185682	2	AC008732	AC008732 Homo sapi	803	19	0.7	578	11	G56129	G56129 SHGC-101597

804	19	0.7	582	10	MUSKR52FP	l28802 Mouse zinc	877	19	0.7	4539	9	HSEPH14	AF101168 Homo sapi
805	19	0.7	599	6	AX062600	AX062600 Sequence	878	19	0.7	5849	10	MMGLI2	X99104 M.musculus
c 806	19	0.7	711	9	AF327112	AF327112 Homo sapi	879	19	0.7	6041	9	HUMOREFGIB	D31163 Human mRNA
c 807	19	0.7	874	33	AC088480	AC088480 Giardina i	880	19	0.7	6170	9	AB002350	AB002350 Human mrn
808	19	0.7	885	9	HUM2INCERN	L14843 Human zinc	881	19	0.7	6350	10	AF019074	AF019074 Mus muscu
809	19	0.7	957	6	AX060615	AX060615 Sequence	c 882	19	0.7	8964	6	AX251976	AX251976 Sequence
810	19	0.7	1103	33	AC051877	AC051877 Giardina i	c 883	19	0.7	8964	6	AX344368	AX344368 Sequence
811	19	0.7	1253	9	HUMN3S03	UA2352 Human N33 g	c 884	19	0.7	8964	6	AX346436	AX346436 Sequence
c 812	19	0.7	1360	14	AF007415	AF007415 Grapevine	c 885	19	0.7	8964	6	AX349013	AX349013 Sequence
c 813	19	0.7	1561	6	AX347902	AX347902 Sequence	c 886	19	0.7	11953	10	MUSFBN1A	I29454 Mouse fibr
c 814	19	0.7	1591	6	AX347901	AX347901 Sequence	887	19	0.7	12118	8	AF050455	AF050455 Zea mays
c 815	19	0.7	1592	6	AX347885	AX347885 Sequence	c 888	19	0.7	13084	8	AX347080	AX347080 Sequence
816	19	0.7	1600	5	MSSNAB	222756 M.salmoides	889	19	0.7	20137	6	AR148845	AR148845 Sequence
817	19	0.7	1602	10	MMU18084	U18084 Mus musculu	890	19	0.7	20138	6	AR148652	AR148652 Sequence
818	19	0.7	1775	6	AX282656	AX282656 Sequence	891	19	0.7	23071	6	AR148847	AR148847 Sequence
819	19	0.7	1831	9	AK027834	AK027834 Homo sapi	c 892	19	0.7	25342	9	AL157892	AL157892 Human DNA
820	19	0.7	1838	9	HSU95044	U95044 Human zinc	893	19	0.7	28088	9	AL159009	AL159009 Human DNA
821	19	0.7	1850	10	BC023179	BC023179 Mus muscu	894	19	0.7	28505	9	HSBAV75A3	HSBAV75A3
822	19	0.7	1894	6	AX274868	AX274868 Sequence	895	19	0.7	28538	2	AC023735	AC023735 Drosophi
c 823	19	0.7	1905	6	AX083427	AX083427 Sequence	c 896	19	0.7	29759	9	AL133392	AL133392 Human DNA
824	19	0.7	1924	3	AF081809	AF081809 Branchios	897	19	0.7	34874	2	AC014553	AC014553 Homo sapi
825	19	0.7	2002	10	AF045565	AF045565 Mus muscu	898	19	0.7	36259	2	AC007046	AC007046
826	19	0.7	2052	9	AK027057	AK027057 Homo sapi	899	19	0.7	39009	9	AC004232	AC004232 Homo sapi
827	19	0.7	2080	9	SS4641	SS4641 H2F-16-Krup	900	19	0.7	39738	2	AC017962	AC017962
828	19	0.7	2088	10	MUSZFCR	L36315 Mus musculu	901	19	0.7	40969	1	SCD6	SCD6
829	19	0.7	2136	9	BC015858	BC015858 Homo sapi	c 902	19	0.7	41275	3	CE57A10B	CE57A10B
830	19	0.7	2166	9	BC017091	BC017091 Homo sapi	c 903	19	0.7	41565	8	AC003952	AC003952 Arabidops
831	19	0.7	2167	9	AK054888	AK054888 Homo sapi	904	19	0.7	42009	9	AC007565	AC007565 Homo sapi
832	19	0.7	2180	9	AK022360	AK022360 Homo sapi	c 905	19	0.7	42030	9	AC007565	AC007565 Homo sapi
833	19	0.7	2188	9	HSMB0063	AL080125 Homo sapi	c 906	19	0.7	43127	9	AC007229	AC007229 Homo sapi
834	19	0.7	2316	9	AK023094	AK023094 Homo sapi	c 907	19	0.7	43276	9	AC074214	AC074214 Homo sapi
835	19	0.7	2330	9	HSMB02196	AL137711 Homo sapi	c 908	19	0.7	44796	2	AC105183	AC105183 Homo sapi
c 836	19	0.7	2397	9	AK057590	AK057590 Homo sapi	c 909	19	0.7	46197	2	AC104764	AC104764 Homo sapi
837	19	0.7	2398	9	HSRTKEPH	227409 H.sapiens m	c 910	19	0.7	47177	2	AL157391	AL157391 Human DNA
838	19	0.7	2399	10	BC021528	BC021528 Mus muscu	c 911	19	0.7	47782	2	AC068445	AC068445
839	19	0.7	2421	9	AK023456	AK023456 Homo sapi	c 912	19	0.7	47782	2	AL157933	AL157933 Human DNA
840	19	0.7	2433	10	MUS2NF5MK	M6516 Mouse zinc	c 913	19	0.7	54228	2	AC104987	AC104987 Homo sapi
841	19	0.7	2457	9	AF244088	AF244088 Homo sapi	c 914	19	0.7	54764	9	AL354952	AL354952 Human DNA
842	19	0.7	2465	9	BC010996	BC010996 Homo sapi	c 915	19	0.7	55256	9	HSJ735G18	HSJ735G18 Human DNA
843	19	0.7	2504	9	AF274348	AF274348 Homo sapi	c 916	19	0.7	55551	2	AC109309	AC109309 Homo sapi
844	19	0.7	2536	9	AK026980	AK026980 Homo sapi	c 917	19	0.7	55619	2	AC109175	AC109175 Mus muscu
845	19	0.7	2547	6	AX306086	AX306086 Sequence	918	19	0.7	56085	2	AC024477	AC024477 Homo sapi
846	19	0.7	2547	10	MMZT3PROT	Z67747 M.musculus	919	19	0.7	56507	2	AC055770	AC055770 Homo sapi
847	19	0.7	2678	9	HSU66561	U66561 Human krupp	c 920	19	0.7	57000	2	AC069278	AC069278 Homo sapi
848	19	0.7	2699	9	AF274347	AF274347 Homo sapi	c 921	19	0.7	57144	2	AC099929	AC099929 Mus muscu
849	19	0.7	2866	9	AF155101	AF155101 Homo sapi	c 922	19	0.7	57623	2	AC105212	AC105212 Homo sapi
850	19	0.7	2933	9	BC007717	BC007717 Homo sapi	c 923	19	0.7	58070	2	AC013628	AC013628 Homo sapi
851	19	0.7	2934	10	AF131197	AF131197 Mus muscu	c 924	19	0.7	58664	2	AC104011	AC104011 Homo sapi
852	19	0.7	3000	9	AF060503	AF060503 Homo sapi	c 925	19	0.7	58668	2	AC100411	AC100411 Mus muscu
853	19	0.7	3014	10	BC022935	BC022935 Mus muscu	c 926	19	0.7	59000	9	AP003128	AP003128 Homo sapi
854	19	0.7	3046	9	AK024047	AK024047 Homo sapi	c 927	19	0.7	59012	9	HSABLG82	HSABLG82 Human ABL g
855	19	0.7	3095	9	BC022932	BC022932 Homo sapi	c 928	19	0.7	60612	2	AC079212	AC079212 Homo sapi
856	19	0.7	3101	9	AF192913	AF192913 Homo sapi	c 929	19	0.7	60683	2	AL356312	AL356312 Human DNA
857	19	0.7	3240	6	AR148651	AR148651 Sequence	c 930	19	0.7	60756	2	AL583823_3	Continuation (4 of
858	19	0.7	3240	6	DB8827	DB8827 Homo sapien	931	19	0.7	61017	2	AC099990	AC099990 Mus muscu
859	19	0.7	3244	6	AR148649	AR148649 Sequence	932	19	0.7	62192	9	AP003033	AP003033 Homo sapi
860	19	0.7	3264	6	AR148650	AR148650 Sequence	933	19	0.7	62456	2	AC099992	AC099992 Mus muscu
861	19	0.7	3268	6	AR148648	AR148648 Sequence	934	19	0.7	63017	2	AC102215	AC102215 Mus muscu
862	19	0.7	3312	9	BC008805	BC008805 Homo sapi	935	19	0.7	63687	8	AH1300474	AH1300474 Antlrtrin
863	19	0.7	3351	9	AK023751	AK023751 Homo sapi	936	19	0.7	63716	2	AC103806	AC103806 Homo sapi
864	19	0.7	3367	9	HSA276182	AJ276182 Homo sapi	937	19	0.7	65608	6	AC107337	AC107337 Homo sapi
865	19	0.7	3370	6	AX335780	AX335780 Sequence	938	19	0.7	65608	6	AX332442	AX332442 Sequence
866	19	0.7	3370	6	HUMTKR	M18391 Human tyros	939	19	0.7	65608	6	AX335496	AX335496 Sequence
867	19	0.7	3469	9	BC014267	BC014267 Homo sapi	940	19	0.7	65608	9	HSU62293	HSU62293 Human LIM-k
868	19	0.7	3580	9	HSMB02544	AL162031 Homo sapi	941	19	0.7	65647	2	AC103937	AC103937 Mus muscu
869	19	0.7	3614	9	AK027712	AK027712 Homo sapi	942	19	0.7	66004	2	AC027365	AC027365 Homo sapi
870	19	0.7	3665	9	BC013382	BC013382 Homo sapi	943	19	0.7	66751	2	AC100850	AC100850 Homo sapi
871	19	0.7	3780	9	HSU25435	U25435 Human trans	944	19	0.7	66980	2	AC013796	AC013796 Homo sapi
872	19	0.7	3782	9	AF385187	AF385187 Homo sapi	945	19	0.7	67059	2	AC103779	AC103779 Homo sapi
873	19	0.7	3810	6	AR081786	AR081786 Sequence	946	19	0.7	67462	2	AC091595	AC091595 Homo sapi
874	19	0.7	3826	9	AF317549	AF317549 Homo sapi	947	19	0.7	67789	9	AC073312	AC073312 Homo sapi
875	19	0.7	4049	9	AF027218	AF027218 Homo sapi	c 948	19	0.7	68334	9	HSJ318B	AL034403 Human DNA
876	19	0.7	4053	9	AF027219	AF027219 Homo sapi	c 949	19	0.7	68509	2	HSY313F4_3	Continuation (4 of

c 950	19	0.7	72044	2	AC026582	Homo sapi
c 951	19	0.7	72045	9	AC092214	Homo sapi
c 952	19	0.7	78167	9	AC092028	Homo sapi
c 953	19	0.7	78396	9	AL590486	Homo sapi
c 954	19	0.7	81017	2	AC025010	Homo sapi
c 955	19	0.7	81990	2	AC024372	Homo sapi
c 956	19	0.7	82251	17	AC004506	Homo sapi
c 957	19	0.7	82453	9	AC004558	Homo sapi
c 958	19	0.7	83051	9	AL611962	Homo sapi
c 959	19	0.7	84096	2	AC074170	Mus muscu
c 960	19	0.7	84113	9	AC073063	Homo sapi
c 961	19	0.7	84613	9	AL136140	Homo sapi
c 962	19	0.7	86348	8	AC006300	Arabidops
c 963	19	0.7	87552	9	AL031771	Homo sapi
c 964	19	0.7	89212	9	AC016679	Homo sapi
c 965	19	0.7	89348	9	AL359734	Homo sapi
c 966	19	0.7	89885	30	HS057333	Homo sapi
c 967	19	0.7	92086	2	AF216668	Homo sapi
c 968	19	0.7	92361	2	AC105766	Homo sapi
c 969	19	0.7	93066	9	HSJ224E15	Homo sapi
c 970	19	0.7	93458	9	AC079344	Homo sapi
c 971	19	0.7	93902	2	AC011090	Homo sapi
c 972	19	0.7	94162	9	AL450322	Homo sapi
c 973	19	0.7	94359	9	AP000974	Homo sapi
c 974	19	0.7	95274	9	HS65019	Homo sapi
c 975	19	0.7	95896	9	AL356384	Homo sapi
c 976	19	0.7	95963	2	AC106422	Rattus no
c 977	19	0.7	98415	9	AC073421	Homo sapi
c 978	19	0.7	98848	2	AC090275	Homo sapi
c 979	19	0.7	99370	9	AC005057	Homo sapi
c 980	19	0.7	100000	9	AB020862	Homo sapi
c 981	19	0.7	100296	9	AL590308	Homo sapi
c 982	19	0.7	100416	9	AC095065	Homo sapi
c 983	19	0.7	100515	9	AL133289	Homo sapi
c 984	19	0.7	100733	9	HS20118	Homo sapi
c 985	19	0.7	101854	9	AC091818	Homo sapi
c 986	19	0.7	102135	2	AF215843	Homo sapi
c 987	19	0.7	102198	9	AL157787	Homo sapi
c 988	19	0.7	103933	9	HS684024	Homo sapi
c 989	19	0.7	104397	2	AF216670	Homo sapi
c 990	19	0.7	104597	9	HS393P12	Homo sapi
c 991	19	0.7	104901	9	AC090423	Homo sapi
c 992	19	0.7	106783	2	AC010455	Homo sapi
c 993	19	0.7	107781	2	AC095175	Rattus no
c 994	19	0.7	108162	2	AC020960	Mus muscu
c 995	19	0.7	108245	9	AC005166	Homo sapi
c 996	19	0.7	108365	2	AC097020	Rattus no
c 997	19	0.7	108503	9	AL627424	Homo sapi
c 998	19	0.7	108634	2	AC020813	Mus muscu
c 999	19	0.7	108907	9	AL135911	Homo sapi
c1000	19	0.7	109056	2	AC074053	Homo sapi

ALIGNMENTS

RESULT	AB013897	2772 bp	mRNA	linear	PRI 03-NOV-1999
LOCUS	AB013897				
DEFINITION	Homo sapiens mRNA for HKR1, partial cds.				
ACCESSION	AB013897.1	GI:6177784			
VERSION	HKR1.				
KEYWORDS	Homo sapiens leukemia cell cell_line:CMK86 CDNA to mRNA.				
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
ORGANISM	Homo sapiens				
REFERENCE	Oguri, T., Katoh, O., Takahashi, T., Isobe, T., Kuramoto, K., Hirata, S.,				
AUTHORS	1 (sites)				
TITLE	The Kruppel-type zinc finger family gene, HKR1, is induced in lung				
JOURNAL	Cancer by exposure to platinum drugs				
	Unpublished (1998)				

REFERENCE 2 (bases 1 to 2772)
 AUTHORS Katoh, O.
 TITLE Direct Submission
 JOURNAL Submitted (13-MAY-1998) Osamu Katoh, Res. Inst. Radiation Biology and Medicine, Hiroshima Univ., Department of Environment and Mutation; Kasumi 1-2-3, Minami-Ku., Hiroshima, Hiroshima 734-8553, Japan (E-mail:katoh@mc.med.hiroshima-u.ac.jp, Tel:+81-82-257-5817, Fax:+81-82-256-7104)

FEATURES
 source
 1..2772
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="19"
 /map="19q12-19q13.1"
 /cell_line="CMK86"
 /cell_type="Leukemia cell"
 1..2096
 /gene="HKR1"

gene
 1..2096
 /gene="HKR1"
 /function="transcription factor"
 /note="The HKR1 protein is similar to Kruppel-type zinc finger transcription factor and has Kruppel-associated box (KRB); A and B box at the N-terminus(nucleotides 193-390) followed by 13 C2H2-type zinc finger domain at the C-terminus(1020-2084)."

CDS

/product="HKR1"
 /protein_id="BA86058.1"
 /db_xref="GI:6177785"
 /translation="GALSMLPGKASLRSPHLCSLPDLRLODPARRAPHLHSETLPFSKSTOETRMATGILRAKKEAFVFRVAVYFQDEERLLSPARTLHREMLETYNHLVLEIPSSKPKLJAOLRGEAPRERKCPDLCPSPKPIOLSPSPCLFSSGOALSOHWLHLSOLFSSLMAGNHLHGKHYPDOKOOPDFCSGAEWIOGGEDSRLRGRVSKNTSKALSPPEEQDPAQKENTYVDISSLERRADLEETKVLHGLEVSGREIKVEEFGPEIKESNLSLQKTYQETPYMTTEWGDSTGSMVLIKNPRTHS GGPVYCRECGRGTFTWKNLITTHORTSGEPYVCKDGRFTWKNLITTHORTSGL KPYVCKECSGFSLSKSLITTHORATGEPYVCKDGRFTWKNLITTHORTSGEPYV YICRECEGFSOKSLIHLRTHNGEKYVCTEGGRHPSMKSLITTHORTSGEPYV CLECGCFSLKSNLKHORSHGEPYVCTEGCGGFTWKNLITTHORTSGEPYVCA ECGRPFNDKSLISHORSHGSEKPEKMECCRRRORKNLIRKRAHSGAVYCECGO GFCARLTILKHORAHAGKPHVCRPCGGGFSRSHLIRHORTSGEKYICRKGGRF SRKSLIRHORTSGC"

polyA-site
 BASE COUNT 746 a 671 c 720 g 635 t
 ORIGIN

Query Match 100.0%; Score 2772; DB 9; Length 2772;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2772; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	caggcgcgtaagctggtggtgagccgggagagcctccttaagcttcccaacc	60
DB	1	CAGGCGCGTTAAGCTGTTGGAGACCGGGAAGCCCTCTTAAGGTTTCCACACC	60
QY	61	tctgctccctgttacctgaacttcgggtcagagatccggcggtgacccggctcattc	120
DB	61	TCTGCTCCTGTTACCTGACTTTCGCTTCAAGATCCCGGCGTGCACCCGCTTCATC	120
QY	121	tgtcttcctgagacttgcctctctccaggaagagactcaaggaagacgaagaatgata	180
DB	121	TGTCTTCTGAGACTTTCGCCCTTCTCCAGGAAGACACACAGACGACGAAATAGGCTA	180
QY	181	caggcgctcctgagagacaaagagagcgcttcggtgacattcaaggatctggtgtact	240
DB	181	CAGGCGCTCCTGAGAGACAAAGAGAGCGCTTCGTGACATTCAGGATGTGGCTGTACT	240
QY	241	tcaccagagagagtggtggtggtgagccctgcagagagacccctgacagggaggtga	300
DB	241	TCACCCAGAGAGAGTGAGGTTGTTGAGCCCTGTCAGAGGACCCCTGCACAGGAGGTGA	300

QY 301 tgcgtgagactataacatctggtctcactggaattccactcttaacaaactca 360
|||||
Db 301 TGCtGAGACTTATACCATGTGCTCAGTGAATTCCATCTTTAAACAAACTCA 360
QY 361 ttgctcagctggaagcggaagcgccctggaagagagaagaalgtccactgacc 420
|||||
Db 361 TTGCTCAGCTGAGCGAGGAGCGCCCTGAGAGAGAGAAATGTCCACTGGAGCC 420
QY 421 tcttccgaatcgaaagcaaaattcaacttagtccctccctgacctgattcttcca 480
|||||
Db 421 TCTGTCGAATTCGAGCGAAGAAATTCACCTTAGTCCCTGCTGTGATTTTCTCA 480
QY 481 gtccgaagctctcagccaacatgtgtgctgtagtcalctctcactgtgttcaagt 540
|||||
Db 481 GTCAGCAAGCTCTCAGCAACATGTGTGAGTCATCTCTCAGCTGTGTTCAAGTT 540
QY 541 tatggcgagaaatccctctcactctggaataaactataccagaaga tcaagaacaagc 600
|||||
Db 541 TATGGCGAGAAATCCTCTCAGCTGGGAAAAACATATCCAGAAATCAGAAACAGC 600
QY 601 aggatccactctgcttagtggaagaagatgatatcaagagagagaaagcctcagac 660
|||||
Db 601 AGGATCCATTTGCTTGTAGTGGCAAGACAGATGATTCAGAGAGGAGAAACTCCAGAC 660
QY 661 tccgttttggagagtaagcaaaaatggtcacttcaaaagcacttccagcccaactgaag 720
|||||
Db 661 TCTGTTTGGAGAGTAAAGCAAAATGSCACTTCAAGGCACTTTCCAGCCCACTGAAG 720
QY 721 aacaaacggccagcagctcccaagagaacaacacagtggtgataataaggtccagcctg 780
|||||
Db 721 AACAAACCCGACAGTCCAGAGAGAACACAGTGTGATATAGGTCCACCCCTG 780
QY 781 aacgagagcgagactaagagaacaagaacagatagtcagtgtttagaattcagat 840
|||||
Db 781 AACGAGGCGAGATCTAGAGAAACAGCAAGATTTGCATGTGTTAAGATCTCAGAT 840
QY 841 ttgagagaatcaaatatagaagatttggcgagccttcaagaagatcaaaactccta 900
*
Db 841 TTGAGAGAAATCAATATGTAAGATTTGGCCAGGCTTTATAGAGATCAAAACCTCTTA 900
QY 901 gctccagaagaacaactggaagacacttacaatgatacgtagtggaagagacact 960
|||||
Db 901 GCTCCAGAAACACAACTGGGAGACACTTACATGTACATGATGAGGAGACACT 960
QY 961 ttggcagatgtcagcttca tcaaaaacccaagagacacactcttgggggaaagcctta 1020
|||||
Db 961 TTGGCAGATGTAGTCTCTCAATCAAAACCAAGGACACACTCTGGGAAAGCTTATG 1020
QY 1021 tgtcagaggaatgtgtgagagccttaccgttgaagtcacaaactgtacacatcagaaga 1080
|||||
Db 1021 TGTGCAAGGAATGTGGGAGGCTTTAGTGGAACTGATTCACACTCAGAGAGA 1080
QY 1081 cacactcagggagaaacttatgtgtgcaagatgtgtgagagagccttacttgaagt 1140
|||||
Db 1081 CACACTCAGGGGAGAAACCTTATGTGTGCAAGATTTGTGGAGAGGCTTTACTTGAAGT 1140
QY 1141 cgaactcttcaacatcaagcgagacacatcagagcttcaagcctta tglgtgcaagaaat 1200
|||||
Db 1141 CGAACCCTTTACACATCAGGAGACACTCAGGCTCAAGCCTTATGTGTCAGAGAT 1200
QY 1201 gttggcagaagctttagcctgaagtcacaaactcatacccaagaagagcgcaacttggg 1260
|||||
Db 1201 GTGGCAGAGCTTAAAGCTGAAGTCAAACTTATTCACACAGAGGGGCGCACACTGGGG 1260
QY 1261 aagaacctatgtttgagaggaatgttggcggtggttccgcaagaatcaacccctgca 1320
|||||
Db 1261 AGAACCTTATGTTTGCAGGGAATGTGGCGTGTGCTTGCCAGCATTCACCTGTGCTA 1320
QY 1321 gacacagaagacacattcagagagaagccttacaatttgcagggaggtgtgcaagct 1380
|||||
Db 1321 GACACAGAAGGACACATTTCAGAGAGAAAGCCTTACATTTGCAGGAGAGTGTGAGCAAGGT 1380
QY 1381 ttagccagaagtcacactcatcagaacacttaaggacacacagagagagagccttatg 1440
|||||

Db 1381 TTAGCCAGAAGTCAACACTCATCAGACACTTAAGAGACACACAGAGAGAAAGCTTATG 1440
QY 1441 tatcagaagaatgtgtgagcgttactttagctgtgaaatacaaaactaaacaacagagga 1500
|||||
Db 1441 TATCAGAGAAATGTGGGAGTCACTTGTAGCTGGAATCAAACTTAAACACACAGAGGA 1500
QY 1501 cacactcaggggtttaaacttatgtctcgtgaggttcggcgcaatgtgtttagcctgaagt 1560
|||||
Db 1501 CACACTCAGGGGTTAAACCTTATGTCTGCTGAGAGTGGGCAATGTCTTAGCTGAAGT 1560
QY 1561 caaaccttaacaaaacacagagtcacacacagcgggggaagcatalgtatgtaaggaat 1620
|||||
Db 1561 CAACCTTAAACAAACACAGAGTCAACACAGGGGGAAGCCATTGTATGTACGAGT 1620
QY 1621 gtggcgagagcttaccgggaatcaacccctgtatcagcaccaagaggaacacacgaagg 1680
|||||
Db 1621 GTGGCGAGGCTTATCCCGGAATCAACCTGTATCCACCAAGAGACACACTCAGGGG 1680
QY 1681 aagaaccttga tglgtcagtgtagtgagcagagccttaatgataagtcacacctcatlt 1740
|||||
Db 1681 AGAAGCCATTTGTATGTCTCAGTGTGACGAGCTTTAATGATATAGTCCACCTCATTT 1740
QY 1741 cacacagaagacacatcagggggaagccttata tglgcagggagtggtgcagaaggt 1800
|||||
Db 1741 CACACAGAGGACACATTTCAGGGGAAAAAGCCTTTATGTGAGGAGAGTGTGCGAGAAAGT 1800
QY 1801 ttggcagaagaacctaaccctgttttaggcagaagagggacactcaggtgtccttltgttca 1860
|||||
Db 1801 TTGCGCAGAAACCTTAACCTGTTTAGCGCAAGAGGGGCAACACTCAGTCTTTGTGTCA 1860
QY 1861 gggagtggtgcagaagccttltgtcctaagtttaactctatataaacaacagagagcacag 1920
|||||
Db 1861 GGGAGTGTGGCAAGGCTTTGTGCTAGTAACTCATTAATTAACACAGAGAGACAGC 1920
QY 1921 caggggggaagcctcactgtgtgcagggagtggtgcagagcctttagccggcagtcacacc 1980
|||||
Db 1921 CAGGGGGAACCTCATGTGTGTCAGAGAGTGTGGCAAGGCTTTAGCCGGAGTCAACCC 1980
QY 1981 tcaatgaacacagaagagacacatcagagagaagaaccttata tltgcagaagtgtagac 2040
|||||
Db 1981 TCATTAAGACACAGAGGACACTTCAGAGAGAAACCTTATATTTGCGAAGAAAGTGTGAC 2040
QY 2041 ggggctttagtcogaagtcacaactatcagacatcagaagacacacacacagatagaagac 2100
|||||
Db 2041 GGGGCTTTAGTGTGGAAGTCCAACTTATCAGACATCAAGAGACACACTCAGGATGAAGC 2100
QY 2101 ttatgtgtataggaatgtgttagagcctttagcagaaggtcacttcaatcagaagcc 2160
|||||
Db 2101 TTTATGTATAGGGAATGTGTACAGCTTTAGCCAGAGATCATATTCATCAGACACC 2160
QY 2161 aagaagacacacagtgctgtgcttltcagcacttgcataagatacaaaagttagaagat 2220
|||||
Db 2161 AGAGACACACACATGCTGCTGCTTTTCAAGCCATTTGCTATGATACCAAGTGGAGACT 2220
QY 2221 tctgtgtgtatgatgatgatgaagctgtacgtglaagacttgatctccatccactgaag 2280
|||||
Db 2221 TCTGTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2280
QY 2281 gagaattgtgtcatttcaagagcctgtcccttccactcactgtgagtggtgtgtgtg 2340
|||||
Db 2281 GAGAAATCTGCTGCTATTTTCAAGAGCCCTGCTTCTCATCTGTGAGTGGGCTTGG 2340
QY 2341 gaaacccgtcagtgatgatgtgcaagagagcagcacaatgcccagcagatagaggt 2400
|||||
Db 2341 GAAACCCGCTCAGTATATGATGTGTGAGAGAGAGAGTCAAAATGCCAGGACATAGGGGT 2400
QY 2401 ggttacctgtglaaaacccaactlaaagctgaagacagtcoccggtctaaatccctcactg 2460
|||||
Db 2401 GGTACTGTGTAAACCAACCTTAAGCTGAAGCAGTCCGCGCTTAATCTCTACTACTG 2460
QY 2461 aattagaacctgtcttccacttgggtgtgtgtcttcccgatgtatcccaaccttacc 2520
|||||

Db 2461 AATTGAAACCTGCTCTCCATTGGTGTGCTTTCCCTCCGATTGATCCAAACCTTCACC 2520
Oy 2521 tatttaagataacgtcccttcccaattggttttatacagtcgtgcccactttga 2580
Db 2521 TATTTACGATACCTGCTTCCTTAATGGTTTATACAGTCGTGTGCCACCTTTTGA 2580
Oy 2581 gtgtgtcccttgataactatacaatcaagcaagtgatattccctatcttgagccataa 2640
Db 2581 GGTGTGCTTTGATCTACTACAAATCAGTCAACGTATATCCCTATCTGAGCCATFA 2640
Oy 2641 aagaccacaactcagctcgtgagagagaagaatacaccctgctgtgagagttggagca 2700
Db 2641 AAGACCAACACTGCTGAGAGAGAGAAATCACCTGCTGTGTGAGAGTTGGAGACA 2700
Oy 2701 ctcctcgtacccctccacacagagagctgtctcttggtcacaataatctcttacc 2760
Db 2701 CTCCCTGCACTCCCTCTCCACAGAGAGCTGTCTTTGTCTCAATAAATCTTTTCTACC 2760
Oy 2761 catctcaccct 2772
Db 2761 CATCTCACCT 2772

RESULT 2

AK056243 2854 bp mRNA linear PRI 31-OCT-2001
LOCUS Homo sapiens CDNA FLJ13681 fis. clone NT2R12005315, highly similar
DEFINITION to Homo sapiens mRNA for HKR1.

ACCESSION AK056243
VERSION AK056243.1 GI:16551590
KEYWORDS oiligo capping; fis (full insert sequence);
SOURCE Homo sapiens teratocarcinoma cell_line:NT2 CDNA to mRNA,
clone.lib:NT2R12 clone:NT2R12005315.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

AUTHORS

1 (sites)
Ishibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S.,
Ono,Y., Holuta,T., Hirooka,S., Murakawa,K., Takiguchi,S.,
Kusano,J., Watanabe,M., Fujimori,K., Tanai,H., Ishida,M.,
Yamashita,H., Chiba,Y., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,
Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y.,
Saito,K., Nishikawa,T., Kimura,K., Matsuo,K., Nakamura,Y.,
Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Takahashi-Fuji,A.,
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
Nagahara,K., Masuko,Y., Nagai,K. and Sugiyama,T.
NEDO human cDNA sequencing project
unpublished

TITLE

JOURNAL

2 (bases 1 to 2854)

Ishigai,T., Otsuki,T. and Sugiyama,T.

Direct Submission

Submitted (24-OCT-2001) Takao Ishigai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kizazuru, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp; Tel:81-48-52-3951, Fax:81-48-52-3952)

NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: RAB and
HRI.

Location/Qualifiers

1..2854

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="NT2R12005315"

/cell_line="NT2"

/cell_type="teratocarcinoma"

/clone_lib="NT2R12"

/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal
precursor cells treated 2-weeks mitotic inhibitor after
5-weeks retinoic acid (RA) induction. -majorly NT2 neuron"

BASE COUNT 761 a 692 c 754 g 647 t
ORIGIN

Query Match 86.0%; Score 2383; DB 9; Length 2854;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2433; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

334 aaattccattcttaaaccaaaacattatgctcagcttgagagagagagcccttga 393
Db 421 AATTTCATCTTCTTAACCAAAATCATTTGCTGAGTGGAGGAGGAGCCCTGGA 480
Oy 394 gagagagagaaatgctccactgagacccctgtgccagaatcgaagcagaatcaacta 453
Db 481 GAGAGAGAGAAATGCTCCACAGACCTGTGCCAATCGAAGCAGAAATTCACCTTA 540
Oy 454 gtccctcgtcccttgatattctccagtcagcaagcttcagccaactgtgtgctga 513
Db 541 GTCCCTCCCTGCCCCCTGATTTCTCCAGTCAGCAAGCTTCAGCCAAATGTGTGCTGA 600
Oy 514 gtccctcgtcccttgatattctccagtcagcaagcttcagccaactgtgtgctga 573
Db 601 GTATCTCTCTCAAGCTGTGTTTCAAGTTATGAGGAGAAATCTCTCCAGCTGGAAAC 660
Oy 574 actatccagaagatcagaacaacacagcagagatccatctgctttagtggaagcagaat 633
Db 661 ACTATCCAGAAATGATCAGAAACAAACAGCAGATCCATTTGCTTTAGTGCAAGCAGAA 720
Oy 634 ggaattcaag 693
Db 721 GGAATTCAGAGGAG 780
Oy 694 caaaggaacttccagcccaacttgagaagaacaacagcagcagcagcagcagcagcag 753
Db 781 CAAAGGCAATTTCCAGCCACCTGGAAGAAACACACACACACACACACACACACACAC 840
Oy 754 cagtggtgatalaaggttcagcccttgaaacagagagagagagagagagagagagag 813
Db 841 CAGGTGTGATATGAGGTCCACCTTGAAACGAGGAGGAGATCTTAAGAAACAGCAAAAG 900
Oy 814 tatttcagtgatttaagaagctcagagatttggaagaatcaataaagaagtttgagcaag 873
Db 901 TATTCATGCTTTAAGAGCTCAGATTTGGAGAAATCAAAATGAGAGGTTTGAGCCAG 960
Oy 874 gcttatacaagagagcaaacccctctagcctcagaagaacacaaacttgagagacact 933
Db 961 GCTTATATCAGAGATCAAAACCTCTTAGCCTCAGAGACACAACTGGGAGACACCTT 1020
Oy 934 acatgtacacgtgagtgaggagagacagcttgagcagatagtcctcatcaaaacccaa 993
Db 1021 ACATGTACACTGAGTGGGAGAGACAGCTTGGCAGATATGCATCCATCAAAACCCAA 1080
Oy 994 ggaacacacttgagggaagagccttctgtgaggaagatgtgggagagccttaagtga 1053
Db 1081 GGACACACTCTGGGGAAGCCCTTATGTGTGAGGGAATGTGGGAGGAGCTTTAGGTGA 1140
Oy 1054 agtcaaacctgtatcacacatcagaagagacacactcaggggagagaaacttatgtgtc 1113
Db 1141 AGTCAAACTGATATCAGATCAGAGACACACTGAGGAGGAGAAACCTTATGTGTGCAAG 1200
Oy 1114 atgtgtgagagagccttacttggaagtcgaaccccttcaacatcagcggagacacacag 1173
Db 1201 AATTGTGAGAGAGGCTTACTTGTGAAGTGACACTTTACATCAGGAGGAGACACTGAG 1260
Oy 1174 ggtctaagcctatgtgtgcaaggaatgttggaagagctttagccttgatgaacccca 1233
Db 1261 GGTCAAGCCTTATGTGTGCAAGGATGTGGCAAGCTTTAGCCTGGAAGTCAAAACCTCA 1320
Oy 1234 ttacacacagagagagacacacttgaggagagagccttatgtttgcagaggaatgtggcgtg 1293
Db 1321 TTACGACACAGAGGAGGAGCAGACACTGGGAGAAACCTTATGTGTGCAAGGAGATGTGGCGTg 1380
Oy 1294 gcttgcgcagcatcacacactgtgtcagacacagaagagacacatcagagagagagcctt 1353

Db	1381	GCTTGGCCAGCA	FTCCAC	ACTG	TGTAGAC	AC	AC	AG	AG	GC	AC	ATTTC	C	AG	AG	A	GA	AG	AG	C	CTT	1440	
Oy	1354	acatttcagggag	tgtgag	caag	gctttag	ccaga	agtcac	acctc	atc	agac	cttaa											1413	
Db	1441	ACATTTCAGG	AGTGTG	AG	CA	AG	CG	CTTTA	AG	CA	AG	CA	CA	CTT	C	AT	C	AG	CA	CTTAA		1500	
Oy	1414	ggacacacagga	ggagaa	aaac	cttat	gtatga	aaaga	atgtg	gggt	taact	ttag	ctgtga										1473	
Db	1501	GGACACAC	AC	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG		1560	
Oy	1474	aatcaaacctc	aaacacac	acacag	aggagac	actcag	ggggtt	aaac	cttat	gtctgtc	gtg											1533	
Db	1561	AATCAAACTC	AAAAC	AC	AC	AC	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG		1620	
Oy	1534	agtgccggag	tgcttgtag	cgctga	agtc	aaacct	taac	aaacac	acag	gggtcc	acac	ggg										1593	
Db	1621	AGTGGGGC	AGTGTCTTT	AG	CTTGA	AG	CTTGA	AG	CTTGA	AG	CTTGA	AG	CTTGA	AG	CTTGA	AG	CTTGA	AG	CTTGA	AG		1680	
Oy	1594	ggggaagac	catlgtat	gtatga	ggag	tgtgtg	ggcgag	cttacc	ccgga	aaatca	aaacctga											1653	
Db	1681	GGGAGAG	AG	CCATT	TGTAT	GTAT	GTAT	GTAT	GTAT	GTAT	GTAT	GTAT	GTAT	GTAT	GTAT	GTAT	GTAT	GTAT	GTAT	GTAT		1740	
Oy	1654	tcacgcacc	agagga	acac	actc	agggg	agaag	ccattgt	atgtc	gtagtgtg	agag											1713	
Db	1741	TCACGCA	CC	AG	AG	AC	AC	CTTCA	CGG	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG		1800	
Oy	1714	gcttaata	gtata	gtccac	ccctcat	ttca	acc	caaga	ggac	atctc	agggg	aaaa	agc	ctt								1773	
Db	1801	GCTTAAT	GAT	AA	GTCC	AC	CCCT	ATTTC	AC	CC	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG		1860	
Oy	1774	tlatgtc	agggag	tgtgtg	agaa	agtlctc	ggca	gaag	cc	tac	ctgtt	atg	ag	ca	aga							1833	
Db	1861	TTATGT	GT	CAG	GG	AG	TGTG	CG	CA	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG		1920	
Oy	1834	ggggaac	acta	agtg	gcctt	gtgtgtg	agagg	atgtgtg	ggga	agc	ctt	gtgt	at	gt	aa	gt	aa					1893	
Db	1921	GGGCA	CA	CT	CA	CT	CA	CT	CA	CT	CA	CT	CA	CT	CA	CT	CA	CT	CA	CT		1980	
Oy	1894	ctccat	ta	aa	ca	ca	ag	ag	ca	ca	ag	gg	gg	gg	gg	gg	gg	gg	gg	gg		1953	
Db	1981	CTCTCAT	TAA	AC	AC	AC	AC	AC	AC	AC	AC	AC	AC	AC	AC	AC	AC	AC	AC	AC		2040	
Oy	1954	ggcaag	c	ctt	ta	g	cc	gg	c	ag	t	c	a	c	c	c	c	c	c	c		2013	
Db	2041	GGCAAG	CG	CTT	TA	GG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG		2100	
Oy	2014	agcc	ta	ta	tt	tg	ca	aa	ag	t	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg		2073	
Db	2101	AGCCT	TAT	ATTT	GC	AG	AA	AG	TG	TG	GA	GG	GG	GG	GG	GG	GG	GG	GG	GG		2160	
Oy	2074	a	tc	ag	a	g	a	c	a	c	a	c	t	c	a	c	t	c	a	c	t		2133
Db	2161	ATCAG	AG	AG	CA	CA	CT	CG	AG	TAG	AA	CTT	AT	GT	AT	GT	AT	GT	AT	GT		2220	
Oy	2134	ggccag	gag	tca	tact	tca	ta	ga	cac	ccaga	ggac	aa	ca	ag	t	g	t	g	t	g		2193	
Db	2221	GGCAG	AG	CTT	AT	CTT	AT	CTT	AT	CTT	AT	CTT	AT	CTT	AT	CTT	AT	CTT	AT	CTT		2280	
Oy	2194	catgtc	ag	ta	aca	aa	ag	tggag	ac	at	ctgtgtg	at	at	at	at	at	at	at	at	at		2253	
Db	2281	CATTG	CT	GT	AG	AC	CA	AA	GT	GG	AG	CA	CT	GT	GT	GT	GT	GT	GT	GT		2340	
Oy	2254	aagact	gt	a	t	c	a	t	c	a	c	c	t	g	a	a	g	a	a	t	c		2313
Db	2341	AAGACT	GT	AT	CT	CA	TT	CA	TT	CA	TT	CA	TT	CA	TT	CA	TT	CA	TT	CA		2400	
Oy	2314	cttcc	ta	ct	a	ct	a	ct	a	ct	a	ct	a	ct	a	ct	a	ct	a	ct		2373	
Db	2401	CTTCT	CT	CA	CT	GT	AG	TG	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT		2460	
Oy	2374	caglt	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a		2433	

Db	2461	CAGTAAATGCCCAGGACAGATGAGGGGTGGGCTACGTGGTGTAAMACCAACCTTAAGCTGAA	2520
Qy	2434	gaacagtccecgactaaatcctcaactgaatgtagaacactgtcttcccaattggtgctt	2493
Db	2521	GACACGTCCCGGCTAAATCCCTCATCTAGTAATGAAACCGCTTCCCATTTGGTGGCTT	2580
Qy	2494	tccctcagatgtatcccaaacctcacaacttttaacgataactgtcccttccctaattgt	2553
Db	2581	TCCTCGAGTATGATCCCAACCCCTTCACACTTTTACGTATACCTGCGCCTTCTCAATTGGT	2640
Qy	2554	tttcaactgctgtgcccacacttttgaagtgtgcttgcactaactaaatcagtaaac	2613
Db	2641	TTTTACACTGCTGTGGCCCACTTTTGAGTGGTCCCTTTCATACCTTACAAATCAGTCAAC	2700
Qy	2614	ggtattccccctatctctgtgcccataaaagacacagacacagctgagtgagagagaa	2673
Db	2701	GTGTATTCCTCCCTATTCTGTAGGCCATTAAGACCCAGACTCAGCTGCAGTGTAGAGAGAAA	2760
Qy	2674	tcaacctgctgtgaggtgtgagggaacactccctgcacttccctccactgaagctgttc	2733
Db	2761	TCACCCTGCTGTGGGGGTTGGGGACACACCTCCGCAATCCCTTCACATGAGAGCTGTTC	2820
Qy	2734	ttttgctcaataaatctcttcttaccactctc	2767
Db	2821	TTTTGCTCAATTAATTCCTTTCTACCCATCCTC	2854
RESULT	3		
LOCUS	AC008806/c	135173 bp	DNA linear PRI 27-APR-2000
DEFINITION	Homo sapiens chromosome 19 clone CTD-2086020, complete sequence.		
ACCESSION	AC008806		
VERSION	AC008806.4	GI:7656695	
KEYWORDS	HTG.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 135173)		
TITLE	DOE Joint Genome Institute and Stanford Human Genome Center.		
JOURNAL	Direct Submission		
REFERENCE	2 (bases 1 to 135173)		
AUTHORS	DOE Joint Genome Institute.		
TITLE	Direct Submission		
JOURNAL	Submitted (03-APR-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE	3 (bases 1 to 135173)		
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-APR-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
REFERENCE	On Apr 27, 2000 this sequence version replaced gi:6600982.		
AUTHORS	Draft Sequence Produced by DOE Joint Genome Institute		
TITLE	www.jgi.doe.gov		
JOURNAL	Finishing Completed at Stanford Human Genome Center		
REFERENCE	www.shgc.stanford.edu		
AUTHORS	Quality: Phrap Quality >=40 99.5% of Sequence;		
TITLE	Estimated Total Number of Errors is 0.6.		
JOURNAL	SHGC-141611 G63592		
REFERENCE	SHGC-37349 G30531		
AUTHORS	SHGC-86030 G61627		
TITLE	Location/Qualifiers		
JOURNAL	1..135173		
REFERENCE	/organism="Homo sapiens"		
AUTHORS	/db_xref="taxon:9606"		
TITLE	/chromosome="19"		
JOURNAL	/clone="CTD-2086020"		
REFERENCE	41590 a 27234 c 27240 g 39109 t		
AUTHORS	ORIGIN		

OY 2587 ccttgcatactacaatcaacagtcgtatctccctatctcgagccataaagacc 2646
|||||
Db 124781 CCTTGCATACCTAGCAATCAAGCAGTGATNTCCCTATCTGTGAGCCCAATAAAGACC 124722
OY 2647 cagactcaagctcagtgagagagaatacacctctgtgtgaggttgaggaccctcct 2706
|||||
Db 124721 CAGACTCAGCTGACGTGAGAGAGAAATCACCCCTGCTGTGGGGTGGGACACATCCCT 124662
OY 2707 gcatccctcctcaagagacgtctcttctgctaataaactcttctcctaccct 2766
|||||
Db 124661 GGATCCCTCTCCACTGAGAGAGCTGTCTTTGCTCAATAAATCTTTCTACCCATCCCT 124602
OY 2767 caccct 2772
|||||
Db 124601 CACCCT 124596

RESULT 4
AC016590 253217 bp DNA linear HTG 20-APR-2001
LOCUS Homo sapiens chromosome 19 clone CTD-3220F14, WORKING DRAFT
DEFINITION
AC016590
AC016590.6 GI:13699590
AC016590
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 253217)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 19
Unpublished
2 (bases 1 to 253217)
DOE Joint Genome Institute.
Direct Submission
Submitted (04-DEC-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Apr 20, 2001 this sequence version replaced gi:7711568.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 940643, BC905667
Center clone name: CITB-EL_3220F14

Summary Statistics
Consensus quality: 207648 bases at least Q40
Consensus quality: 225569 bases at least Q30
Consensus quality: 232428 bases at least Q20
Estimated insert size: 250510; agarose-1p estimation
Estimated insert size: 249417; sum-of-ctligs estimation
Quality coverage: 9.03 in Q20 bases; agarose-1p estimation
Quality coverage: 9.07 in Q20 bases; sum-of-ctligs estimation.
NOTE: This is a 'working draft' sequence. It currently
* consists of 39 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1019: contig of 1019 bp in length
* 1020 1119: gap of unknown length
* 1120 1152: contig of 1033 bp in length
* 2153 2252: gap of unknown length
* 2253 3701: contig of 1449 bp in length
* 3702 3801: gap of unknown length
* 3802 5053: contig of 1252 bp in length
* 5054 5153: gap of unknown length

5154 6184: contig of 1031 bp in length
* 6185 6284: gap of unknown length
* 6285 7809: contig of 1525 bp in length
* 7810 7909: gap of unknown length
* 7910 8986: contig of 1077 bp in length
* 8987 9086: gap of unknown length
* 9087 10090: contig of 1004 bp in length
* 10091 10191: gap of unknown length
* 10191 11225: contig of 1035 bp in length
* 11226 11326: gap of unknown length
* 11326 12365: contig of 1040 bp in length
* 12366 12465: gap of unknown length
* 12466 14535: contig of 2070 bp in length
* 14536 14635: gap of unknown length
* 14636 16489: contig of 1854 bp in length
* 16490 16589: gap of unknown length
* 16590 19109: contig of 2520 bp in length
* 19110 19209: gap of unknown length
* 19210 21841: contig of 2632 bp in length
* 21842 21941: gap of unknown length
* 21942 23735: contig of 1794 bp in length
* 23736 23835: gap of unknown length
* 23836 25861: contig of 2026 bp in length
* 25862 25961: gap of unknown length
* 25962 29117: contig of 3156 bp in length
* 29118 31907: gap of unknown length
* 31907 32007: contig of 2690 bp in length
* 32008 36364: gap of unknown length
* 36364 36464: contig of 4337 bp in length
* 36465 36464: gap of unknown length
* 36465 43309: contig of 6845 bp in length
* 43310 43409: gap of unknown length
* 43410 48185: contig of 4776 bp in length
* 48186 48285: gap of unknown length
* 48286 55626: contig of 7341 bp in length
* 55627 55726: gap of unknown length
* 55727 65781: contig of 10055 bp in length
* 65782 65881: gap of unknown length
* 65882 66673: contig of 3792 bp in length
* 66674 69773: gap of unknown length
* 69774 79389: contig of 9616 bp in length
* 79390 79489: gap of unknown length
* 79490 91453: contig of 11964 bp in length
* 91454 91553: gap of unknown length
* 91554 96671: contig of 5118 bp in length
* 96672 96771: gap of unknown length
* 96772 100544: contig of 3773 bp in length
* 100545 100644: gap of unknown length
* 100645 111423: contig of 10779 bp in length
* 111424 111523: gap of unknown length
* 111524 115278: contig of 3755 bp in length
* 115279 115379: gap of unknown length
* 115380 124011: contig of 8633 bp in length
* 124012 124111: gap of unknown length
* 124112 138475: contig of 14364 bp in length
* 138476 138575: gap of unknown length
* 138576 140613: contig of 2038 bp in length
* 140614 144949: gap of unknown length
* 144950 145049: contig of 4236 bp in length
* 145050 164352: contig of 19303 bp in length
* 164353 164452: gap of unknown length
* 164453 170540: contig of 6088 bp in length
* 170541 170640: gap of unknown length
* 170641 185046: contig of 14406 bp in length
* 185047 185146: gap of unknown length
* 185147 222209: contig of 37063 bp in length
* 222210 253217: gap of unknown length
* 253217 253217: contig of 30908 bp in length.
Location/Qualifiers
1.253217
/organism="Homo sapiens"
/db_xref="taxon:9606"

/chromosome="19"
/clone="CFD-3220F14"
/clone_1lb="Caltech human BAC library D"
BASE COUNT 62246 a 62867 c 62532 g 61752 t 3820 others
ORIGIN

Query Match 81.0%; Score 2244; DB 2; Length 253217;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2344; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 427 cagaatcgaagccgaagaattcaacttagtccctccctccctcgtatcttccagtcagc 486
|||||
Db 74511 CAGAAATGGAAGCCAGAAATTCACCTTAGTCCCTCCCTCCCTATTTTCCAGTACG 74570

QY 487 aagctcccaagcccaacatctgtctgagtcactctctcagctgtttcgaagttacgg 546
|||||
Db 74571 AAGCTCCAGCCACATGTGTGGCTGAGTCATCTCTCACTCTTTTCAAGTTTATGGC 74630

QY 547 caagaaatccctccctccctcgtggaacacactccagaagatcagaacaaacagcagatc 606
|||||
Db 74631 CAGGAATTCCTCTCCACTGGGAATACATATCCAGAAATCAGAAACACAGCAGATC 74690

QY 607 catcttccttaagtggaagcagaatggaatctcaagagagagagagactccaagctctgt 666
|||||
Db 74691 CATTCCTCTTAGTGGCAAGCAGAAATGATTCAGAGGGAAGACTCCAGACTCTGT 74750

QY 667 ctggagagatgaagcaaaatctgacatcaagagcacttccagcccaactgaagaacaac 726
|||||
Db 74751 TTGGGAGAGTAAGCAAAATATGCGACTTCMAAGCAGCTTTCAGCCCACTGMAAACAAC 74810

QY 727 agccagcagctcccaaggaagacaacagctgtgtatagataggtccagccctgaacgga 786
|||||
Db 74811 AGCCAGCAGATCCACAGGAAGACACACAGTGTGATATAGGGTCCAGCCCTTAAAGGA 74870

QY 787 gggcagaatcagaaggaacagacaagattgtcagtggttaagaagttccagatcttgag 846
|||||
Db 74871 GGGCAGATCTTAGAGGAAGACAGCAAAAGTATTTGCAATGTTAAGAGTCCAGGATTTGGAG 74930

QY 847 aatcaaatatgaagagcttgggccaagcttatacaagagatcaaacctccctgaagctcc 906
|||||
Db 74931 AAATCAATATGAAGATTTGGGCGCAGCTTTATCAAGAGAGTCAAACTCTTATGCTCC 74990

QY 907 aagaagacaaactgggagagacacttacaatgtacactgagtgaggagagacagcttgga 966
|||||
Db 74991 AGAAGACACAAACTGGGGAACACCTTACATGTACACTGAGTGGGAGACAGCTTTGGCA 75050

QY 967 gtagtgaagctcactcaaaaacccaaggaacacactctggggggaagccttatgtgtgca 1026
|||||
Db 75051 GTATGTCAGTCTCATCAAAAACCCAAAGACACACTCTGGGGGAAGCCTTATGTGTGCA 75110

QY 1027 gggaaatctgggcaagccttcaagtggaagtcacaaactgtatcacacatccagaagacacac 1086
|||||
Db 75111 GGGAAATGCGCGCAGAGCTTTACGTGAAGTCAAACTGTATCACACATCAGAGACACACT 75170

QY 1087 caaggaggaacaaacttatgtgtgcaaggaatgtggaagagagccttacttgaagtgcgaac 1146
|||||
Db 75171 CAGGAGGAACCTTATGTGTGCAAGGATGTGAGCAGAGCTTACTTGAAGTCGAAC 75230

QY 1147 tctttacacatcaagcagacacactcaaggtccaagccttatgtgtgcaaggaatgtgtggc 1206
|||||
Db 75231 TCTTTACATCAAGCGGACACACTCAGGGCTCAAGCCTTATGTGTGCAAGGAATGTGGC 75290

QY 1207 aagagcttaagcgaagatcaaacacttataccacagagagagcagacatgggagagagc 1266
|||||
Db 75291 AGAGCTTTAGCCTTAGACTCAAACTCTTACCCACAGAGGCGCACACTGGGAGAGAGC 75350

QY 1267 ctaatacttgcaggaagatgtggcgctgtgcttccgacagacatcaacacatgtgtcagaaca 1326
|||||
Db 75351 CTTATGTTTCAGAGGATGTGGGCTGTGTCGCCAGCATTCACACTGTGTGTCAGACACA 75410

QY 1327 agagagacacatcaagagagagagccttacaatttgcaggaagtgtagcaagcctttagcc 1386
|||||

Db 75411 AGAGGACACATTCAGAGAGAGAGCCTTACTTTGACAGGAGTGTGACCAAGCCTTATGCC 75470

QY 1387 agaagtcacacccatcaagaacacttaagagacacacagagagagaacacttatgtatgca 1446
|||||

Db 75471 AGAGTACACACTCATATAGACACTTAAAGACACACAGAGAGAGAACTTATGTATGCA 75530

QY 1447 cagaatctggcgctcaactttagcttggaatcaaacctcaaaaacacacagagagacacac 1506
|||||
Db 75531 CAGAAATGGGCGCTACCTTAGCTGGAATCAAACTCAAAACACACACAGAGACACACT 75590

QY 1507 caagggttaaaccttatgtctgtcgtggaatgtggggagagtgctttagcctgaagtcaaac 1566
|||||
Db 75591 CAGGGGTTAAACCTTATGTGTGCTGGAGTCCGGCGAGTCTTTAGCTTAAAGTCAAAAC 75650

QY 1567 ttaacaacacacagagagtcacacacagggggaagagacatttgtatgaagagtggtggc 1626
|||||
Db 75651 TTAAACAAACACAGAGTACACACAGGGGGAAGAACCATTTGTATGTACGAGTGTGGC 75710

QY 1627 gaagcttaccggaatcaaccctgtatcaagcagagagacacactcagggggaagc 1686
|||||
Db 75711 GAGCTTTACCCGGAATCAACCTTGACACGACACAGAGACACACTCAGGGGAGAGC 75770

QY 1687 catcttgatctgtcgtgagtggtgagcagagcttataatgaatccaccttatccaacc 1746
|||||
Db 75771 CATTTGATGTGTGAGTGTGACAGAGCTTTAATGATATGATCCACTTATTTACACACC 75830

QY 1747 agagagacacatcaagggaaagaccccttatactgtgcaagagtggtgcaagagttcggc 1806
|||||
Db 75831 AGAGGACACATTCAGGGGAAGAACCTTTATGTGCAAGAGAGTGTGCAAGAGTTTCGC 75890

QY 1807 aagaagcctaacctgtttagcacaagagaggaacactcaagtgcttctgtgtgcaagagat 1866
|||||
Db 75891 AGAAGCCTAACCTGTTTAGCAGACAGAGGCGCACACTAGTGTGCTTTGTGCAAGGAGT 75950

QY 1867 gtaggcaagcctttagtgcctaaatctcaatcaaacacagagagagagagagagag 1926
|||||
Db 75951 GTGGCAAGGCTTTTGTGCTAAAGTAACTCTCAATTAACACCAAGAGACACAGCGGG 76010

QY 1927 ggaagcctcatctgtgtcagagagtggtggaagagcttlaacggagtcacacccaata 1986
|||||
Db 76011 GGAAGCTCATGTGTGAGGAGAGTGTGGCAAGGCTTTATCCCGGACATCACTCATTA 76070

QY 1987 gacaacagagagacacatcagagagagaagccttatatttgcagaagtggtggaagagct 2046
|||||
Db 76071 GACACACAGAGACACATTCAGAGAGAGAACCTTATATTTCAAAAGTGTGAGCAGGGGCT 76130

QY 2047 ttagtgcgaagctcaaaccttatcaagacatcaagagagagagagagagagagagagag 2106
|||||
Db 76131 TTATGTGGAAGTCAACCTTATCAAGACATCAAGAGACACACTCAGAGATTAAGAACTTTATG 76190

QY 2107 tgtataggaatgtgtgtacagcctttagcagagagtcatacttcatcaagacacagaaga 2166
|||||
Db 76191 TGTATAGGGAATGTGTGACAGCCTTTAGCCAGGAGTCAATCAATCAGACACAGAGGA 76250

QY 2167 cacaacagtgctgtgtgcttctcaagcattgtctagatlacaaagtgtgagacatctgtgt 2226
|||||
Db 76251 CACACACAGTGTGTGCTTTTACAGCATGTCTAGATACCAAGGTGAGACACTTCTGTG 76310

QY 2227 tgtgattatgcatgagagagtgactgtgtaagacttgtatcccatccacccgaagagagat 2286
|||||
Db 76311 TGTGATTTATGATGAGAGCTGTACTGTGTAAGCTTGTATCTCCACTCACTGAAGGGAAT 76370

QY 2287 tgcctgcctcatcttcaagagccctgcctcccaactgtgtgaatgtgtgtgtgtggaac 2346
|||||
Db 76371 TGCTGGCTCATTTTTCAGAGAGCCTTGCCCTTCTCACTGTGTGATGTGTGTGTGTGAAAC 76430

QY 2347 cgttcaagtaatgaatgtggaagagagtgcaaatgtcccaagcagataggggtgtgtac 2406
|||||
Db 76431 CGGTCAAGGTATGATAGTGTGAGAGGAGGAGTCAAAATGCCAGGACAGTATAGGGTGTGATC 76490

QY 2407 ctgtgtaaacccaacacttaagctgaagagagatccgggtcctaaatcctcatatcgaattga 2466
|||||
Db 76491 CTGTGTAACCAACCTTAAAGCTGAAGACAGTCCGGGCTAAATCTCTCATACATGAATTTGA 76550

OY 2467 gaactgtctcccaattggtgtgtcttcctcgaattgatcccaacccttaccatttt 2526
|||||
Db 76551 GAACCTGCTTCCCATTTGGTGTGCTTCCTCCGATGATCCCAACCCCTCCACTATTTT 76610
OY 2527 acgtatactgcctcttccctaattgttttatactgctgtgcccacttttgaatgtg 2586
Db 76611 ACGTATACCTGCGCTTCTTCTATTTGTTTACACTGCTGTGSCCACTTTTGAATGGTG 76670
OY 2587 ccttgcatactataaatcagtaagtaagttatctccctattctgaagccataaaagacc 2646
|||||
Db 76671 CTTTGCACTACTTACAAATCACTCAACGATGATTCCTCTATTTCTAGCCCATMAAAGACC 76730
OY 2647 cgaactcagctcagtggaagaagaatcaaccctgtctgtgaagttggaagaccactccct 2706
|||||
Db 76731 CAGACTCAAGCTCAGTAGAGAGAGAAATCACCTGCTGTGGGGGTTGGGGACACTCCCT 76790
OY 2707 gcaaccctctcactgaagagctgtcttcttgcataataaattcttcttaccatcct 2766
|||||
Db 76791 GCATCCCTCTCCACTGAGAGCTGTCTTTGCTCAATMAATCTTTTCTACCATCCT 76850
OY 2767 caccct 2772
|||||
Db 76851 CACCT 76856
|||||
RESULT 5
AC093235 38173 bp DNA linear PRI 29-SEP-2001
LOCUS AC093235
DEFINITION Homo sapiens chromosome 19 clone L1NLR-267A4, complete sequence.
ACCESSION AC093235
VERSION AC093235.2 GI:15808549
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 38173)
DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE
AUTHORS 2 (bases 1 to 38173)
DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (16-AUG-2001) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 38173)
DOE Joint Genome Institute and Stanford Human Genome Center.
AUTHORS Direct Submission
TITLE Direct Submission
JOURNAL Submitted (29-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Sep 29, 2001 this sequence version replaced gi:15193369.
COMMENT Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Flinishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.
FEATURES
source
1. 38173
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="L1NLR-267A4"
BASE COUNT 9495 a 8465 c 8118 g 12095 t
ORIGIN
Query Match 79.1%; Score 2193; DB 9; Length 38173;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2343; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 427 cagaatcgagccagaataactagatccctcctcctccttgatttctccagtcagc 486

|||||
Db 28888 CAGATTCAGAGCAAAATCAACTTAGTCCCTCCCTGCTCTGATTTTCTCCAGTCAGC 28947
OY 487 aagccttaagcacaatgltgtgctgagtcacatctctcagctgttctcaagttatgg 546
Db 28948 AAGCTTCAGCCAAACATGTGTGGCTGAGTCATCTCTCAGCTGTTTCAAGTTATGGG 29007
OY 547 caggaaacccctccacccttggaagaacacatccagaagatcagaagaacagcagatc 606
|||||
Db 29008 CAGGAATCTCTTCCACTGGGAAAACATATCCAGAAATCAGAAACACGACGATC 29067
OY 607 caticgtcttagttagtgcagaagcaatgatatcaagaggaggaagcctccagctccgt 666
Db 29068 CATCTGCTTTAGTGGCAAGACAGAAATGATTCAGAGGAGAGAACTCCAGACTCTGT 29127
OY 667 ttggagagtgtaagcaaaaatgtgcaacttcaagagcacttccagccactgaagaacac 726
Db 29128 TTGGAGAGTAAAGCAAAATGGCACTTCAAAAGCACTTCCAGCCCACTGAAGAACAC 29187
OY 727 agccagcagctccaaaggaaagacacacagtggtgatataaggtccagccctgaagga 786
Db 29188 AGCCAGCAGTCCAGAGAGACACACATGTGATATAGGCTCCAGCCCTGAACGGA 29247
OY 787 gggcagatctagaggaacacagacaagatltgcatgtgttgaagltctcaagattggag 846
Db 29248 GGGCAGATCTAGAGGAACAGACAAAGTATTCATGTGTTAGAGTCTCAGGATTTGGAG 29307
OY 847 aatcaataatagaagaatttggccagagcttatacaagaggtcaaacctcccttaacc 906
Db 29308 AAATCAAAATATGAAAGATTTGGGCCAGGCTTTATCAAGAGTCAAAACCTCTTACCCTCC 29367
OY 907 aagaagacaacactgggagacaccttaccatgtacactgaagtgggagagacacttgca 966
Db 29368 AGAAGACACAAACTGGGAGACACCTTACATGTACACTGAGTGGGAGACACTTTGGCA 29427
OY 967 gtagtcaagltctcaataaaacccaaggaacacacttgggggaaagccttatgtgtca 1026
Db 29428 GTATGTCAAGTCTCTCATCAAAACCCAGAGACACACTCTGGGGAAAGCCTTATGTGTGA 29487
OY 1027 gggaaatgtggcgaagcctttagtgaagttcaaacctgaatcaacatcagaagaacact 1086
Db 29488 GGGAAATGTGGCGAGGCTTTACGTGAGATCAAACTGATCAACATCAGAGACACACT 29547
OY 1087 caggagagaaaccttatgtgtgcaagagatgtgacagagcttacttggaaagtcgaacc 1146
Db 29548 CAGGGGAGAAACCTTATGTGTCAAGAGATTGGACGAGCCTTTACTTGGAAAGTGAAC 29607
OY 1147 tctttacacatcagcgagacacatcagggtcacaagccttatgtgtgcaaggaattgggc 1206
Db 29608 TCTTTACACATCAGCGGACACACTCAGGGCTCAAGCCTTATGTGTGCAAGGAATGTGGGC 29667
OY 1207 aagaacttagcctgaagttcaaacctcatcaccacaagggcgagacacttggggaagac 1266
Db 29668 AAGAGCTTAGCTTGAGGTAAACTCTATTACCCACAGGGGCGCACATGGGGGAAMC 29727
OY 1267 ctatgttgcaggaatgtggcgtgtgcttcgcagagatccacacactgtgtcagaaca 1326
Db 29728 CTTATGTTGACGAGGAAATGTGGCGGTTCGCGAGATTCACACCTGTGTGACAGACA 29787
OY 1327 aagaagacacatcaggaagaagccttaccatttgcagggagtgtgagcaagccttagcc 1386
Db 29788 AAGAGCACATTCAGAGAGAGAGCCCTTACATTGTGACGGGAGGTGTGAGCAAGCCTTAGCC 29847
OY 1387 aagaagtcacactcatcagaacacttaagaacacacagagagaagagccttatgtatga 1446
Db 29848 AGAAGTCACACTCATCAGACACTTAAGGACACACAGAGAGAGACCTTATGTATGCA 29907
OY 1447 cagaatgttggtcactttagcttggaatcaaacctcaaaaacacaccagagagacact 1506
Db 29908 CAGAAATGTGGCATACCTTAGCTGGAATCAAACTCAAAACACACCGAGAGACACT 29967
OY 1507 cagggttaaaccttatgtctgtcctgaggtgcggagcagtgcttagcctgaagtaaac 1566
|||||

Db 29968 CAGGGGTTAAACCTTATGCTGCTGAGAGTGGGGCAGTCTTTAGCCTGAAGTCAACCC 30027
 Qy 1567 ttaacaacacacagaggttcaacacaggggagagccatttgaatgaagagtgatgagc 1626
 Db 30028 TTTAAACAAACCAAGAGTTCACACACGGGGGAGAAAGCCATTGTTATGACGAGTGTGGC 30087
 Qy 1627 gaagctttacccggaatcaacccctgaacacacagagagacactcaaggaggaagc 1686
 Db 30088 GAAGCTTTACCCGGAATTCACACCTTGACACGACACAGAGACACACTCAGAGGGAGAAC 30147
 Qy 1687 catttgaatgctgaagtgatgagcagagctttaaagtaagtcaccccttaattcaacc 1746
 Db 30148 CATTGATGATGCTGAGTGTGAGAGAGGCTTTAATGATTAAGTCCACCTCAATTTTCACACC 30207
 Qy 1747 agagagacacattcaaggagaaagccctttaaagtcaaggagtgatgagagaggttcagc 1806
 Db 30208 AGAGGACACATTCAGGGGAAAGCCTTTATGTGSCAAGGAGTGTGCGAAGAGTTTCGGC 30267
 Qy 1807 agaagcctaacctgtttaggcacacagagagacactcaagtgcccttgatgagagagat 1866
 Db 30268 AGAAGCTTAACCTGTTAGGCACAGAGGGCACACTCAGTCCCTTGTGTGACAGGAGT 30327
 Qy 1867 gtagcgaagcctttgtgctaaagtaattcctcaatlaaacccagagagacacagcagggg 1926
 Db 30328 GTGGGCAAGGCTTTGTGCTAAGTTAACTCTCAATTAACACAGAGAGCACACGACAGGG 30387
 Qy 1927 ggaagcctcaatgtcgaagagagtgatgagcagagcttlaacggcagtcacacccctta 1986
 Db 30388 GGAAGCCTCATGTGTGCGAGGAGTGTGGCAGAGCTTTACCGGCACTCACCCTCATTTA 30447
 Qy 1987 gacacagagagacacattcaaggagagagcctataatttgacagaaagtgtgagcagggc 2046
 Db 30448 GACACACAGAGACACATTCAGGAGAGAGCCTTATATTTCAAAAGTGTGAGCAGGGCT 30507
 Qy 2047 ttatgtcgaagtgcaaccccttatagacatlaagagagacacactaagataaagactatg 2106
 Db 30508 TTATGTGGAAGTCAACCTTATCAGACATCAGAGACACACTCAGGATTAACCTTTATG 30567
 Qy 2107 tgatagaggaatgttgatcacagcctttagccagagagatcattcatcacagacacagaga 2166
 Db 30568 TGATAGGGAATGTGGTACAGCCTTTAGCCAGAGATCATCTTCATCAGACACAGAGGA 30627
 Qy 2167 cacacacagtgctgtgcttttcaagccattgctagaatacaaaagtggagacattctgtg 2226
 Db 30628 CACACACAGTGTGCTGCTTTTACAGCATGCTAGATACCAAGTGAAGACATTTCTGTG 30687
 Qy 2227 tggatattatgacagagactgtactgttaagaacttgaatccatccacccgaaggagaaat 2286
 Db 30688 TGTGATTAATGATGACAGACTGTACTGTGAAGACTGTGATTCATCCACCTGAAGAGAAAT 30747
 Qy 2287 tgcgtgctcaatttcagagagcccttgcctcctcactgttgatgagtggtgtggaacc 2346
 Db 30748 TGCTGCTCATTTTTCAGAGACCCCTGCCCTTCCTACATGTGGAATGGTGGTGAACCC 30807
 Qy 2347 cggatcaggttaatgatagtgagcagagagcagtaaatgcccagagcagataaggggtggtac 2406
 Db 30808 CGGTCAAGTAATGATGATGAGCAGAGAGCAGTCAAAATGCCAGGACAGATAGGGGTGGGTAC 30867
 Qy 2407 ctggttaaaccccaacttlaaagtgaagaagctccggcgtcctaaatcctcactactgaattga 2466
 Db 30868 CTGGTAAACCCCAACTTAAAGCTGAAGACAGTCCGGGCTAAATCTCTCACTAGTAATGA 30927
 Qy 2467 gaacctgtctcccatcttggtgtgcttctcctccgattgaatccaaaccttcaactaattt 2526
 Db 30928 GAACCTGTCTCCCATTTGGTGTGCTTCCTCCGATTGATCCCAACCCCTCACCTAATTTT 30987
 Qy 2527 aggtataactgccccttctctaattggtttttaaactgctgtgcccaccccttgaagtgtg 2586
 Db 30988 AGCTAATACCTGCCCTTCCATTAATGGTTTTCACAGCTGCTGGCCACCTTTTGAAGTGCTG 31047
 Qy 2587 ccttgatcaactataaactcaagtcacagctgtatctccctatcttgagccataaagaagac 2646
 Db 31048 CCTTTCATATCTTACAAATCAGTCAACGTTGATTTCCCTATTTCTGAGCCCATTAAGAGACC 31107

Qy 2647 cagactcagctcagtgagagagaaatcacccctgtgtgaggttgaggaccactccct 2706
 Db 31108 CAGACTAGCTGCAGTACAGAGAAATCACCTGCTGCTGGGGTTGGGACCACTCCCT 31167
 Qy 2707 gcalcccccctcactgaagagctgttcttcttgctcaataaattcttcttaccactct 2766
 Db 31168 GCATCCCTCTCCACGTGAGCTGTCTTCTTCTCATATAAATCTTCTTACCCTATCT 31227
 Qy 2767 caccct 2772
 Db 31228 CACCCT 31233

RESULT 6
 BC004513 1998 bp mRNA linear PRI 12-JUL-2001
 LOCUS Homo sapiens, clone IMAGE:3928207, mRNA, partial cds.
 DEFINITION BC004513
 ACCESSION BC004513
 VERSION BC004513.1 GI:13325426
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1998)
 Strausberg, R.
 Direct Submission
 Submitted (12-MAR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 CONTACT: MGC help desk
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: DCTD/DP
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Institute for Systems Biology
 http://www.systemsbio.org
 contact: amadnesystemsbiology.org
 Anup Madan, Rachel Dickhoff, Jessica Fahay, Stephanie Ford, Julia
 Greene, Mark Kellerman and Anuradha Madan

REMARK
 COMMENT
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
 Series: IRAL Plate: 14 Row: b Column: 10
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

FEATURES
 source
 Location/Qualifiers
 1..1998
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3928207"
 /tissue_type="Ovary, adenocarcinoma"
 /clone_lib="NIH-MGC_9"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 1..1280
 /codon_start=3
 /product="unknown (protein for IMAGE:3928207)"
 /protein_id="AAH04513.1"
 /db_xref="GI:13325427"
 /translation="HEVSVSGGEIKYEFPGPIKESNLSTLOKTONGEPYMTW
 GDSFGSMVLIKNRTHSGKRPVYRREGGRFTWKSNIITHQRTHSGKRPVYCDGGR
 GFTWKSNIITHQRTHSGKRPVYCKEGRGFSLSKSLITHQRHHTGEKPYVCEGRGFG
 ROHSHLVHKKRTHSGKRPYICRECEGRGFSLSKSLITHQRHHTGEKPYVCEGRGFG
 KSNLKTQRTHSGKRPVYCLCEGQCFSLKSNLKHQRSHHTGEKPYVCEGRGFRPKS
 TLTHQRTHSGKRPVYCEGGRGFSLSKSLITHQRTHSGKRPYVCEGRGFRPKS
 FRHKAHSGATVCEGGRGFSLSKSLITHQRHHTGEKPYVCEGRGFRPKS
 RTHSGKRPYICRKGGRGFSRKSNIITHQRTHSG"

BASE COUNT 575 a 458 c 510 g 455 t

ORIGIN

Query Match 63.9%; Score 1771; DB 9; Length 1998;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1941; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

826 tagaagctccagagatttggaagaatcaaatatgaagaagcttggccagagcttatcaag 885
10 TAGAGTCTCAGGATTTGGAGAAATCAATATGAAGAGTTTGGCCAGGCTTATCAAG 69
QY 886 agtcaaacctccttagcctccagaagacacaaactlgyggagacactlcatlgtacatg 945
DB 70 AGTCAAACTCCTTACCTCCAGAAACACAACTGGGAGACACTTACATGTACACTG 129
QY 946 aatggggaagacagcttggcagatgatcagctcctcaaaaaaaccaagaagacactc 1005
DB 130 AGTGGGAGACAGCTTGGAGATGTAGTCACTCTCATCAAAAAACCAAGGACACTCTG 189
QY 1006 ggggaagacctatgtgtcagggaaatgtgagcgaagcttcaagtgagtcgaactcga 1065
DB 190 GGGGAAGGCTTATGTGTCTCAGGAGATGTGGGAGAGCTTACGTGGAAGTCAAACTGA 249
QY 1066 tcaacacatcagagagacacatcagaggagaacacttatgtgtgaaagatgtgagcag 1125
DB 250 TCACACATCAGAGACACACTCAGGGAGAAACCTTATGTGTGCAAGGATTTGTGAGCAG 309
QY 1126 gcttacttggaagtcgaacctcttcaacatcagcagacacactcagagctcaaacctt 1185
DB 310 GCTTACTTGGAACTGGAACCTTTTACATCAGCGGACACCTAGGGCTCAAACTCAACCTT 369
QY 1186 atgtgtcgaaggaatgttggcagagactttagcctgaagtcgaacctcatlaccacaga 1245
DB 370 AGTGTGCAAGGATGTGGGAGAGCTTATGAGTGAAGTCAAACTCATTTACCACAG 429
QY 1246 gggcgacaccttgggagaagccttatgttgcagggaaatgttggcgttgcctccagc 1305
DB 430 GGGCGACACTGGGGAAGCCTTATGTTCAGGGAAATGTGGCGCTGGCTTTCGGCAGC 489
QY 1306 attcacactgttcaagacaaagagacacatcagagagaagccttcaatttgcagg 1365
DB 490 ATTCAACCTGCTGACAGACACAAAGAGACACATTCAGAGAGAAAGCTTATTCAGAGG 549
QY 1366 agtgtgaagcaagctttagccagaagtcacacctcatcagacactlaagagacacaga 1425
DB 550 AGTGTAGCAAGGCTTATGAGCAGAGTCAACCTCATCAAGACACTTAAGAGACACACAG 609
QY 1426 gagagaagccttatgtatgcacagaatgttggcgttcaactttagcttggaaataacctca 1485
DB 610 GAGAGAAGCCTTATGTATGACACAGAAATGTGGGCTCACTTACCTGAAATCAAACTCA 669
QY 1486 aaacacacagagagacactcaggggtttaaacttatgtctgcctggagttggcgact 1545
DB 670 AAACACACAGAGAGACACTCAGGGGTTTAACCTTATGTCTGCTGAGCTGGGAGG 729
QY 1546 gctttagcctgaagtcacaaacttaacacacagagatcacacagggggagaagccat 1605
DB 730 GCTTATGCGTGAAGTCAAACTTAACAAACACAGAGGTACACAGGGGAGAAAGCAT 789
QY 1606 tttgtatgaagagatgttggcgaagcttaccgggaatacaacctgtatcagcaccaga 1665
DB 790 TTTGATGTACGGAGTGTGGGCGAGGCTTATCCGGAATCAACCCGTATCACCACACGA 849
QY 1666 ggaacacacagggagaagacacttgtatgttgcagtgatgttgaagagagctttaatgata 1725
DB 850 GGAACACACTCAGGGGAGAAAGCCATTGTATGTCTGCTGAGTGGACGAGGCTTTAATGATA 909
QY 1726 agtcaacctcatlccacacagagagacacatcagggagaagcctttaatgtgcagg 1785
DB 910 AGTCCACCCCTCATTTCCACACAGAGACACATTCAGGGGAAAGCCTTTATGTGACAGG 969
QY 1786 agtgtggcagaaggttctggcagaagccttaacctgtttaggcacaagaggcactcag 1845

DB 970 AGTGTGGCAGAAAGCTTTCGGCAGAAAGCCTTAACCTGTTTATGAGCAAGAGGACACTCAG 1029
QY 1846 gtgccttltgtgcagggagatgttggcaagccttltgttctaagttaactcatcaaac 1905
DB 1030 GTGCCCTTGTGTGACGGAGGTGGGCAAGGCTTTGTCTAAGTTAACTTCATTAAC 1089
QY 1906 accaagagacacagcagagggggaagcctcatgtgtgcagggagatgttggcgaagcctta 1965
DB 1090 ACCAGAGACACACGAGGGGGGAAGCCTCATGTGTGCAAGGAGTGTGGCAAGGCTTAA 1149
QY 1966 gccggcagtcacacctcatltagacaccagaagacacatcagaagagaagccttattt 2025
DB 1150 GCCGGCAGTCACACCTCATTTAGACACAGAGACACATTCAGAGAGAAAGCCTTATATT 1209
QY 2026 ggaagaaggtgtgagggccttaagtgcgaagtcacaccttcaagatacagaagaagac 2085
DB 1210 GCAGAAAGTGTGACGGGCTTATGTGGAAGTCAACCTTATCAACATTCAGAGACAC 1269
QY 2086 actcagatagaactltagtltatagggaaatgttgaagcctttagccagagatcat 2145
DB 1270 ACTCAGATAGAAACTTATGTGTATAGGGAATGTGTACAGCCTTACGACAGAGTCAAT 1329
QY 2146 acttcaacagacacacagagacacacagtgctgtgccttlttcaagcattgtatgata 2205
DB 1330 ACTTCATCAGACACACAGAGACACACAGTGTGTGCTTTTTCAGCCATTGCTAGATA 1389
QY 2206 ccaaaatgagacatctgtgtgtatgtatgtacatgagactgtgtaagactgtatc 2265
DB 1390 CCMAATGAGACATTTCTGTGTGTATGTGATGAGACTGTACTGTAAAGACTTATTC 1449
QY 2266 tccatccacctgaagagaatgtgtgtcatatlttcaagagcccttgccttccatcgt 2325
DB 1450 TCCATCCACCTGAAGAGAAATTCCTGCTCATTTTCAGAGGCCCTCCCTCCTCACTGT 1509
QY 2326 gtagtgggtgtgttggaaacccggatcaggttaagatgagtgaggaggaagtgaaatggc 2385
DB 1510 GGTATGGTGGGTGTGT-6AAACCCGGTCAAGTAATGATGTGCGAGGAGCGTCAAAATGCC 1568
QY 2386 caagcagatagagggtgtgttaccctgtgtgaaccccaaccttaaaagttaagaacagtc 2445
DB 1569 CAGCGAGATAGGGGTGTGATGCTGTGTGAACCCACACTTAAGCTGAAGACATGCCGG 1628
QY 2446 taaatccatactgaattgagaacctgttcccatlttgggtgtgtccttccctcagatga 2505
DB 1629 TAAATCCTCATTAAGTAATGAGAACTGTCTCCATTTGGTGTGCTTCCCTCCGATTGA 1688
QY 2506 tcccaacctcaaccttattagatlaactgtcccttcccttaattgtttttaaactgtc 2565
DB 1689 TCCCAACCCCTTCACTATTTTACGTATACCTGCGCTTCTTAATGTGTTTAACTGTCT 1748
QY 2566 gtgccacacctttagtgtgtccttgcataactcaaatcaatcaagtgatattccctc 2625
DB 1749 GTGCCACCTTTGAGTGTGTGCTTGTGCATTAACATCAATCAAGCGTATTCCTCC 1808
QY 2626 attctgaagccataaaagaccagactcagctgtcagtgagagagaatacaacctgtcgt 2685
DB 1809 ATTCTGAGCCCATTAAGAAAGACCCAGACTCACTCAGTGAAGAGAAATCAACCTGCTGT 1868
QY 2686 ggaaggttggggaacctccctgtcatccctctcaactggaagatgttcttggctaata 2745
DB 1869 GGGGGTGGGAGCACCTCCCTGATCCCTTCACACTGAAGAGCTGTCTTGTCTCAATA 1928
QY 2746 aaattccttctacacactca 2768
DB 1929 AAATTCCTTCTACCATCTCA 1951

RESULT 7
HUMHR1 HUMHR1 672 bp DNA linear PRI 27-APR-1993
LOCUS Human Kruppel related gene, exon X, clone pKHR18.
DEFINITION M20675.M19429
ACCESSION M20675.1 GI:184108
VERSION

KEYWORDS Kruppel-related protein.
SOURCE Human DNA, clone pKRIIRS.
ORGANISM Homo sapiens
REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo. 1 (bases 1 to 672)
AUTHORS Ruppert, J.M., Kinsler, K.W., Wong, A.J., Bigner, S.H., Kao, F.-T., Law, M.L., Senanez, H.N., O'Brien, S.J. and Vogelstein, B.
TITLE The Gli-Kruppel family of human genes
JOURNAL Mol. Cell. Biol. 8, 3104-3113 (1988)
MEDLINE 89096896
COMMENT computer-readable sequence for [1] kindly provided by J.M.Ruppert, 20-JUN-1988.
FEATURES
source location/Qualifiers
1..672
/organism="Homo sapiens"
/db_xref="taxon:9606"
exon <1..>672
/note="exon X"
BASE COUNT 195 a 155 c 184 g 138 t
ORIGIN Chromosome 19.
Query Match 24.28; Score 672; DB 9; Length 672;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 672; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 942 actgaagtcggagagacacatttgagcatatgctcctcctcaaaacccaagagacac 1001
DB 1 ACTGATGGGGAGACACCTTTGGCAGATGTGCTACCTCCATCAAAACCCAGAGACAC 60
QY 1002 tctggggaaagccttatgtgtgcaaggaaatgtggcgaagccttaagtcgaac 1061
DB 61 TCTGGGGAAAGCCTTATGTGTGCAAGGAATGTGGGAGGCTTACGTGAATCAAC 120
QY 1062 ctgatcacatcagagagacactcaggggagaacacctatgtgtgcaagatgtgga 1121
DB 121 CTGATCACACTCAGAGACACACTCAGGGGAGAACTTATGTGTGCAAGATGTGGA 180
QY 1122 cgaagcttacttggaagtcgaacctcttcaacatcagcgagacacacacacacac 1181
DB 181 CGAGGCTTACTTGGAAAGTCGAACCTTTTACATCAGCGGACACCTCAGGCTCAAG 240
QY 1182 ccttatgtgtgcaagaaatgtggcagagctttagcctgaagtcacaaacctatcac 1241
DB 241 CCTTATGTGTGCAAGAAATGTGGCAGACGCTTATGCTGAAGTCGAACCTCATTACCCAC 300
QY 1242 cagagggcgacacactgtggggaagccttatgtgtgcaaggaaatgtggcgttgcgc 1301
DB 301 CAGAGGGCGACACTGTGGGAGAACTTATGTGTGCAAGGAATGTGGGCGCTTTCGC 360
QY 1302 cagcatcacacctgtgcaacacagaagagacacattcagagagaagacattatgac 1361
DB 361 CAGCATTCACACCTGTGCAACACAGACACATTCAGAGAGAAAGCCTTACATTTCG 420
QY 1362 agggagtgtagaagcctttagccagaagtacacacctcaccagacattaaagacac 1421
DB 421 AGGAGTGTGAGCAAGGCTTATGAGCAGAAAGTGTGGGCGTCACTTACCTGAAATCAAC 480
QY 1422 aacggagagaagccttatgtgtgcaagaaatgtggcgtaacctttagctggaatcaac 1481
DB 481 ACAGGAGAGAAGCCTTATGTATGACAGAAATGTGGGCGTCACTTACCTGAAATCAAC 540
QY 1482 ctcaaaacacacagagagacacactcagggtttaaaccctatgtctgctggagtgagg 1541
DB 541 CTCAAAACACACAGAGAGACACTCAGGGGTTTAAACCTTATGTCTCTGGAGATGGGG 600
QY 1542 caatgtccttagcctgaagtcacaaaccttaacacacacagaggttcacacaggggagaa 1601
DB 601 CAGTGTCTTGTAGCTGAAGTCGAACCTTAAACACACAGAGGTTCACACAGGGGGAAG 660
QY 1602 caattgtatgt 1613
|||||

DB 661 CCATTGTATGT 672
RESULT 8
LOCUS AX067361 678 bp DNA linear PAT 24-JAN-2001
DEFINITION Sequence 65 from patent WO0078960.
ACCESSION AX067361
VERSION AX067361.1 GI:12544985
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo. 1 (bases 1 to 678)
AUTHORS Vignli, J. and Mitcham, J.L.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer
JOURNAL Patent: WO 0078960-A 65 28-DEC-2000;
CORIXA CORPORATION (US)
FEATURES
source location/Qualifiers
1..678
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature 1..678
/note="n = A,T,C or G"
BASE COUNT 165 a 167 c 160 g 182 t 4 others
ORIGIN
Query Match 20.6%; Score 571; DB 6; Length 678;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2078 gagagacactcagagatagaacttatgtgtataggaatgtgtacagccttagaca 2137
DB 9 GAGGACACACTCAGAGATAGAACCTTATGTGTATAGGGAATGTGTACACCTTTACCCA 68
QY 2138 ggaagtcacttcatcagacacacagagagacacacagtcgtgtgttttcaagcatt 2197
DB 69 GGAATCTACTTCTCATCGACACACAGAGACACACAGTCTGTGGCTTTTCAACCAT 128
QY 2198 gctagataccagaatgtgagacattctgtgtgtatattgcatagagactgtactgtlaa 2257
DB 129 GCTAGATACCAGAAAGTGAGACATTCTGTGTGTATGTATGATGAGACTGTACGTGAGA 188
QY 2258 ctgtatctccatccacactgaagagagaatgtcgtgcatttccagagacccgtcccttc 2317
DB 189 CTTGTATCTCCATTCACCTGAAGAGAAATGTGGCTCATTTTTCAGGAGCCCTGCTTC 248
QY 2318 ctcaactgtgatatgtgtgtgtgaaacccgylcaaglaaagaaacactaaagctgaagaca 2377
DB 249 CTCATCTGTGATGTGTGTGTGTGAAACCCGGTCAAGTATGATAGTGAAGAGAGAGT 308
QY 2378 caaatgcccagagacagaaatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2437
DB 309 CAATGCCCCAGGACAGATAGGGGTGAGTACTGTGTAACCAACCACTTAAAGCTGAAGACA 368
QY 2438 gtcccggttaataatccatcatatcgaatgtgaacactgtctccatgtgtgtgtgtgt 2497
DB 369 GTCCCGGTAAATTCCTATCTGAATTTGAACCTGTCTTCCATTTGGTGTGCTCTTCC 428
QY 2498 ccgattatcccaacaccttaacattatctgataacctgccccttccatgaattgtttt 2557
DB 429 CCGATTGATCCCAACCTTACCTATTTTACGATACCTGCTTTCCTTAATGGTTTTT 488
QY 2558 acaactgtgtccacactttagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2617
DB 489 ACACTGTGTGCCACCTTTTGAAGTGTGCTTGTGATATTCAATCAATGATCAAGTGT 548
QY 2618 atccacctatctgagccataaagaacca 2648
|||||
DB 549 ATTCCCTATTTCGAGCCCATTAAGACCA 579

RESULT 9
AC024911 130351 bp DNA linear HTG 02-MAR-2000
LOCUS Homo sapiens chromosome 3p clone RP11-420a23, WORKING DRAFT
DEFINITION AC024911
SEQUENCE, 32 unordered pieces.
ACCESSION AC024911
VERSION AC024911.1 GI:7143406
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 130351)
Lu G., Zhao Y., Lin W., Dong H., Wan M., Xu S., Gu W., Tu Y.,
Jia J., Wu C., Zhang C., Zhong M., Zhou Y., Ren S., Fu G., Chen Z.
and Huang M.
Direct Submission
Submitted (02-MAR-2000) Genomic Dept., Chinese National Human
Genome Center at Shanghai, 315 Guo Shou Jing Road, Shanghai,
Shanghai 201203, P. R. China
-----Genome Center Information-----
Center: Chinese National Human Genome Center at Shanghai
JOURNAL
TITLE
REFERENCE
AUTHORS
COMMENT
Web site: <http://www.chgc.sh.cn>
Email: mhuang@chgc.sh.cn or fugang@chgc.sh.cn
-----End Genome Center Information-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 2267: contig of 2267 bp in length
* 2268 4378: contig of 2111 bp in length
* gap of unknown length
* 4379 6650: contig of 2272 bp in length
* gap of unknown length
* 6651 8791: contig of 2141 bp in length
* gap of unknown length
* 8792 11482: contig of 2691 bp in length
* gap of unknown length
* 11483 13872: contig of 2390 bp in length
* gap of unknown length
* 13873 16745: contig of 2873 bp in length
* gap of unknown length
* 16746 18965: contig of 2220 bp in length
* gap of unknown length
* 18966 21604: contig of 2639 bp in length
* gap of unknown length
* 21605 24066: contig of 2462 bp in length
* gap of unknown length
* 24067 26444: contig of 2378 bp in length
* gap of unknown length
* 26445 29243: contig of 2799 bp in length
* gap of unknown length
* 29244 31995: contig of 2752 bp in length
* gap of unknown length
* 31996 34538: contig of 2543 bp in length
* gap of unknown length
* 34539 37970: contig of 3432 bp in length
* gap of unknown length
* 37971 41593: contig of 3623 bp in length
* gap of unknown length
* 41594 44802: contig of 3209 bp in length
* gap of unknown length
* 44803 48691: contig of 3889 bp in length
* gap of unknown length
* 48692 52259: contig of 3568 bp in length

* 52260 55732: contig of 3473 bp in length
* gap of unknown length
* 55733 57756: contig of 2024 bp in length
* gap of unknown length
* 57757 61389: contig of 3633 bp in length
* gap of unknown length
* 61390 65398: contig of 4009 bp in length
* gap of unknown length
* 65399 69740: contig of 4342 bp in length
* gap of unknown length
* 69741 74577: contig of 4837 bp in length
* gap of unknown length
* 74578 80165: contig of 5588 bp in length
* gap of unknown length
* 80166 86055: contig of 5890 bp in length
* gap of unknown length
* 86056 89215: contig of 3160 bp in length
* gap of unknown length
* 89216 94479: contig of 5264 bp in length
* gap of unknown length
* 94480 103136: contig of 8657 bp in length
* gap of unknown length
* 103137 113254: contig of 10118 bp in length
* gap of unknown length
* 113255 130351: contig of 17097 bp in length.
* Location/Qualifiers
1. 130351
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3p"
/clone="RP11-420a23"
BASE COUNT 38785 a 26050 c 26300 g 39176 t 40 others
ORIGIN
FEATURES
source
Query Match 1.3%; Score 37; DB 2; Length 130351;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 37; Conservative 0; Mismatches 0; Gaps 0;
OY 2462 attgagaacctgtctccattgtgtgtcttccccc 2498
Db 69153 ATTGAGACCTGTCTCCATTGTGTGTCTTCCCTC 69189

RESULT 10
AC096898 161625 bp DNA linear HTG 12-JAN-2002
LOCUS Homo sapiens chromosome 4 clone RP11-618K19, WORKING DRAFT
DEFINITION AC096898
SEQUENCE, 3 unordered pieces.
ACCESSION AC096898
VERSION AC096898.5 GI:18139542
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 161625)
Waterston, R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 161625)
Waterston, R.H.
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (01-OCT-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jan 12, 2002 this sequence version replaced gi:17921273.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>

Contact: submissions@watson.wustl.edu
 ----- Project Information -----
 Center project name: H_NH0618K19
 ----- Summary Statistics -----
 Sequencing vector: M13, 0%
 Sequencing vector: plasmid: 100%
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap, version 0.990319
 Consensus quality: 160310 bases at least Q40
 Consensus quality: 160620 bases at least Q30
 Consensus quality: 160738 bases at least Q20
 Insert size: 161425; sum-of-contigs
 Quality coverage: 7.76 in Q20 bases; sum-of-contigs
 Quality coverage: 8.11 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 3023: contig of 3023 bp in length
 * 3024 3123: gap of unknown length
 * 3124 115853: contig of 112730 bp in length
 * 115854 115953: gap of unknown length
 * 115954 161625: contig of 45672 bp in length.
 Location/Qualifiers
 1. 161625
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="4"
 /clone="RP11-618K19"
 1. 3023
 /note="assembly_name:Contig24"
 clone_end:SP6
 vector_side:left
 3124. 115853
 /note="assembly_name:Contig26"
 clone_end:T7
 vector_side:right
 115954. 161625
 /note="assembly_name:Contig25"
 BASE COUNT 47603 a 31872 c 31765 g 50185 t 200 others
 ORIGIN
 Query Match 1.3%; Score 37; DB 2; Length 161625;
 Best Local Similarity 100.0%; Pred. No. 2.4e-09;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2462 atcgagaacctgtccattcgtggtgcttcctc 2498
 ||||||||||||||||||||||||||||||||||||
 DB 110301 ATTGAGAACCTGCTTCCTTCATTTGCTCTTCCTC 110337
 RESULT 11
 AC034197 164264 bp DNA 11near HNG 29-MAY-2000
 LOCUS Homo sapiens chromosome 3 clone RP11-72801 map 3p, WORKING DRAFT
 DEFINITION
 SEQUENCE, 14 unordered pieces.
 AC034197
 AC034197.3 GI:8101273
 HTG: HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 164264)

AUTHORS
 Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,
 Dong,W., Fan,H., Feng,X., Guan,Q., Gu,X., Guo,D., He,L., Hu,S.,
 Huang,F., Jin,Y., Kang,N., Li,C., Li,G., Li,J., Li,L.,
 Li,S., Li,T., Liu,B., Liu,N., Liu,Y., Li,W., Li,Y.,
 Luo,J., Niu,Y., Qi,Q., Qi,X., Song,S., Sun,M., Sun,W., Sun,Y.,
 Tao,R., Wang,H., Wang,J., Wang,J., Wang,L., Wang,L., Wang,R.,
 Wang,X., Wang,X., Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y.,
 Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H.,
 Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y.,
 Zhang,Z., Zhu,B., Yu,J. and Yang,H.
 Chromosome 3p genomic sequence
 Unpublished
 2 (bases 1 to 164264)
 Zeng,Y., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B.,
 Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y.,
 Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y.,
 Li,G., Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D.,
 Huang,F., Zhang,G., Li,D., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.
 and Yang,H.
 Direct Submission
 Submitted (05-APR-2000) Human Genomic Center, Institute of
 Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
 100101, P.R.China
 On May 29, 2000 this sequence version replaced g1:7644466.
 -----Genome Center
 Center:Beijing Center
 Center code:Beijing
 Website:http://hgc.igtp.ac.cn
 http://www.genomics.org.cn
 Contact:hgc@igtp.ac.cn
 ----- Project Information
 Center project name:1% project
 Center clone name: RP11-72801
 ----- Summary Statistics
 Sequencing vector: pUC18; 100% of reads
 Chemistry: Dye-terminator: ET 5% of reads
 Chemistry: Dye-terminator Big Dye: 45% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 15901 bases at least Q40
 Consensus quality: 163151 bases at least Q30
 Consensus quality: 165125 bases at least Q20
 Insert size: 158755; sum-of-contigs
 Quality coverage: 4.57x in Q20 bases;sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 14 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 2043: contig of 2043 bp in length
 * 2044 2143: gap of unknown length
 * 2144 4407: contig of 2264 bp in length
 * 4408 4507: gap of unknown length
 * 4508 5590: contig of 1083 bp in length
 * 5591 8894: gap of unknown length
 * 8894 8994: contig of 3204 bp in length
 * 8995 12938: gap of unknown length
 * 12939 13038: contig of 3944 bp in length
 * 13039 19151: gap of unknown length
 * 19152 24784: contig of 6113 bp in length
 * 24785 24884: gap of 5533 bp in length
 * 24885 32377: contig of 7493 bp in length
 * 32378 32478: gap of unknown length
 * 32479 41545: contig of 9068 bp in length
 * 41546 41645: gap of unknown length
 * 41646 50878: contig of 9233 bp in length
 * 50879 50978: gap of unknown length
 * 50979 62812: contig of 11834 bp in length

FEATURES

source

1. 164264
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p"
/clone="RP11-72801"
1. 2043
/note="assembly_name:Contig4"
2144. 4407
/note="assembly_name:Contig5"
4508. 5590
/note="assembly_name:Contig6
clone_end:77
vector_side:right"
5691. 8894
/note="assembly_name:Contig7"
8995. 12938
/note="assembly_name:Contig8"
13039. 19151
/note="assembly_name:Contig9"
19252. 24784
/note="assembly_name:Contig10"
24885. 32377
/note="assembly_name:Contig11"
32478. 41545
/note="assembly_name:Contig12"
41646. 50878
/note="assembly_name:Contig13"
50979. 62812
/note="assembly_name:Contig14"
62913. 88788
/note="assembly_name:Contig15"
88889. 115986
/note="assembly_name:Contig16"
116087. 164264
/note="assembly_name:Contig17"
46659 a 32302 c 32891 g 51080 t 1332 others

BASE COUNT

ORIGIN

Query Match 1.3%; Score 37; DB 2; Length 164264;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2462 attgagaacctgtctccattgtgtgtgtcttc 2498
|||||

Db 55166 ATTGAGAACCTGCTTCCTCCATTGCTGCTTCCTC 55202

RESULT 12

AC069265

LOCUS

DEFINITION

AC069265 165379 bp DNA linear HTG 29-MAY-2000
Homo sapiens chromosome 3 clone RP11-208K18 map 3p, WORKING DRAFT

SEQUENCE

AC069265

VERSION

AC069265.2 GI:81011151

KEYWORDS

HTG: HTGS_PHASE1; HTGS_DRAFT.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 165379)
Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,
Dong,W., Fan,H., Feng,X., Guan,Q., Gu,X., Guo,D., He,L., Hu,S.,
Huang,F., Jin,Y., Kang,N., Li,C., Li,G., Li,J., Li,L.,
Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W., Li,W., Li,Y.,

TITLE

JOURNAL

REFERENCE

AUTHORS

Lu,J., Liu,Y., Qi,Q., Qi,X., Song,S., Sun,M., Sun,W., Sun,Y.,
Tao,R., Wang,H., Wang,J., Wang,J., Wang,L., Wang,L., Wang,R.,
Wang,X., Wang,X., Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y.,
Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H.,
Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y.,
Zhang,Z., Zhu,B., Yu,J., and Yang,H.

Chromosome 3p genomic sequence

Unpublished

2 (bases 1 to 165379)

Wang,L., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B.,
Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y.,
Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y.,
Li,G., Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D.,
Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.,
and Yang,H.

Direct Submission

Submitted (23-MAY-2000) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing,
100101, P.R.China

On May 29, 2000 this sequence version replaced g1:8039668.

-----Genome Center

Center:Beijing Center

Center code:Beijing

Website:http://hgci.igtp.ac.cn

http://www.genomics.org.cn

Contact:hgci@igtp.ac.cn

----- Project Information

Center project name:1# project

Center clone name:RP11-208K18

----- Summary Statistics

Sequencing vector: pUC18; 100% of reads

Chemistry: Dye-terminator; ET 55% of reads

Chemistry: Dye-terminator Big Dye; 45% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 164602 bases at least Q40

Consensus quality: 165885 bases at least Q30

Consensus quality: 166968 bases at least Q20

Insert size: 162389; sum-of-contigs

Quality coverage: 6.82x in Q20 bases;sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1874: contig of 1874 bp in length

1875 1974: gap of unknown length

1975 8441: contig of 6467 bp in length

8442 8541: gap of unknown length

8542 20983: contig of 12442 bp in length

20984 21083: gap of unknown length

21084 38280: contig of 17207 bp in length

38291 38390: gap of unknown length

38391 56760: contig of 18370 bp in length

56761 56860: gap of unknown length

56861 93526: contig of 36566 bp in length

93527 93627: gap of unknown length

93627 165379: contig of 71753 bp in length.

Location/Qualifiers

1. 165379
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p"
/clone="RP11-208K18"
1. 1874
/note="assembly_name:Contig3"
1975. 8441
/note="assembly_name:Contig4"
8542. 20983

```
misc_feature      /note="assembly_name:Contig5"
                  21084..38290
misc_feature      /note="assembly_name:Contig6"
                  38391..56760
misc_feature      /note="assembly_name:Contig7"
                  56861..93526
misc_feature      /note="assembly_name:Contig8"
                  93627..165379
                  /note="assembly_name:Contig9"
BASE COUNT      51415 a 32065 c 32897 g 48394 t      608 others
ORIGIN

Query Match      1.3%; Score 37; DB 2; Length 165379;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2462 atgagagacctgtctccattgtgtgtcttcctc 2498
      |||
Db 83850 ATTGAGAACTGCTTCCATTGCTGTCTTCCTC 83886

RESULT 13
AC026219      184375 bp      DNA      linear      HTG 29-MAY-2000
LOCUS      Homo sapiens chromosome 3 clone RP11-815J19 map 3p, WORKING DRAFT
DEFINITION      AC026219
ACCESSION      AC026219
VERSION      AC026219.2 GI:8101287
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 184375)
  Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,
  Dong,W., Fan,H., Feng,X., Guan,Q., Gu,X., Guo,D., He,L., Hu,S.,
  Huang,F., Jin,Y., Kang,N., Li,C., Li,G., Li,J., Li,L.,
  Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W., Li,Y.,
  Luo,J., Niu,Y., Qi,Q., Qi,X., Song,S., Sun,M., Sun,W., Sun,Y.,
  Tao,R., Wang,H., Wang,J., Wang,J., Wang,L., Wang,R.,
  Wang,X., Wang,X., Wang,Y., Wu,D., Xu,Q., Xie,F., Xuan,Z., Xue,Y.,
  Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H.,
  Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y., Zhang,Y.,
  Zhang,Z., Zhu,B., Yu,J. and Yang,H.
  Chromosome 3p genomic sequence
  Unpublished
2 (bases 1 to 184375)
  Kang,N., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B.,
  Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y.,
  Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y.,
  Li,G., Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D.,
  Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.
  and Yang,H.
  Direct Submission
  Submitted (21-MAR-2000) Human Genomic Center, Institute of
  Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
  100101, P. R.China
  On May 29, 2000 this sequence version replaced gi:7272036.
  -----Genome Center
  Center:Beijing Center
  Center code:Beijing
  Website:http://hgsc.jgtp.ac.cn
  http://www.genomics.org.cn
  Contact:hgc@jgtp.ac.cn
  -----Project Information
  Center project name:1% project
  Center clone name: RP11-815J19
  -----Summary Statistics
  Sequencing vector: pUC18: 100% of reads
  Chemistry: Dye-terminator: ET 55% of reads
  Chemistry: Dye-terminator: Big Dye: 45% of reads
  Assembly program: Phrap: version 0.990329
```

```
Consensus quality: 146985 bases at least Q40
Consensus quality: 167323 bases at least Q30
Consensus quality: 182494 bases at least Q20
Insert size: 122192, sum-of-coverage
Quality coverage: 3.83x In Q20 bases; sum-of-coverage
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 56 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1      1389: contig of 1389 bp in length
*      1489: gap of unknown length
*      1490: contig of 2113 bp in length
*      3603: gap of unknown length
*      3703: contig of 2445 bp in length
*      6148: gap of unknown length
*      6248: contig of 1594 bp in length
*      7842: gap of unknown length
*      7942: contig of 1628 bp in length
*      9570: gap of unknown length
*      9670: contig of 2301 bp in length
*      11971: gap of unknown length
*      11971: contig of 1139 bp in length
*      12071: 13209: contig of 1139 bp in length
*      13210: gap of unknown length
*      13310: contig of 1397 bp in length
*      14906: contig of 1397 bp in length
*      14907: gap of unknown length
*      15007: contig of 1909 bp in length
*      16916: 17015: gap of unknown length
*      17016: 18655: contig of 1640 bp in length
*      18656: gap of unknown length
*      18756: 20974: contig of 2219 bp in length
*      20975: gap of unknown length
*      21075: 22409: contig of 1335 bp in length
*      22410: 22509: gap of unknown length
*      22510: 24337: contig of 1828 bp in length
*      24338: 24437: gap of unknown length
*      24438: 25826: contig of 1389 bp in length
*      25827: gap of unknown length
*      25926: gap of unknown length
*      27769: contig of 1843 bp in length
*      27770: 27869: gap of unknown length
*      27870: 29441: contig of 1572 bp in length
*      29442: 29541: gap of unknown length
*      29442: 32494: contig of 2953 bp in length
*      32495: gap of unknown length
*      32595: 34633: contig of 2039 bp in length
*      34634: 34733: gap of unknown length
*      34734: 37159: contig of 2426 bp in length
*      37160: 37259: gap of unknown length
*      37260: 39230: contig of 1971 bp in length
*      39231: gap of unknown length
*      39330: gap of unknown length
*      40642: contig of 1312 bp in length
*      40643: 40742: gap of unknown length
*      40743: 42739: contig of 1997 bp in length
*      42740: 42839: gap of unknown length
*      42840: 45245: contig of 2406 bp in length
*      45246: 45345: gap of unknown length
*      45346: 47083: contig of 1738 bp in length
*      47084: 47183: gap of unknown length
*      47184: 49945: contig of 2762 bp in length
*      49946: gap of unknown length
*      50046: 51544: contig of 1439 bp in length
*      51545: 51644: gap of unknown length
*      51645: 53891: contig of 2247 bp in length
*      53892: 53991: gap of unknown length
*      53992: 57221: contig of 3230 bp in length
*      57222: 57321: gap of unknown length
*      57322: 60016: contig of 2695 bp in length
*      60017: 60116: gap of unknown length
*      60117: 62563: contig of 2447 bp in length
```

```

* 62564 62663: gap of unknown length
* 62664 64926: contig of 2263 bp in length
* 64926 65026: gap of unknown length
* 65027 68373: contig of 3347 bp in length
* 68374 68473: gap of unknown length
* 68474 71700: contig of 3227 bp in length
* 71701 71800: gap of unknown length
* 71801 74355: contig of 2555 bp in length
* 74356 74456: gap of unknown length
* 74456 77267: contig of 2811 bp in length
* 77267 77367: gap of unknown length
* 77367 80106: contig of 2740 bp in length
* 80107 80206: gap of unknown length
* 80207 82763: contig of 2557 bp in length
* 82764 82863: gap of unknown length
* 82864 85878: contig of 3015 bp in length
* 85879 85978: gap of unknown length
* 85979 88748: contig of 2770 bp in length
* 88749 88848: gap of unknown length
* 88849 93166: contig of 4318 bp in length
* 93167 93266: gap of unknown length
* 93267 96634: contig of 3368 bp in length
* 96635 96734: gap of unknown length
* 96735 10049: contig of 3715 bp in length
* 10049 100549: gap of unknown length
* 100550 103083: contig of 2534 bp in length
* 103084 103183: gap of unknown length
* 103184 107029: contig of 3846 bp in length
* 107030 107129: gap of unknown length
* 107130 110324: contig of 3195 bp in length
* 110325 110424: gap of unknown length
* 110425 114667: contig of 4043 bp in length
* 114668 114567: gap of unknown length
* 114568 119345: contig of 4778 bp in length
* 119346 119445: gap of unknown length
* 119446 122439: contig of 2994 bp in length
* 122440 122539: gap of unknown length
* 122540 127598: contig of 5059 bp in length
* 127599 127698: gap of unknown length
* 127699 133792: contig of 6094 bp in length
* 133793 133892: gap of unknown length
* 133893 140225: contig of 6333 bp in length
* 140226 140325: gap of unknown length
* 140326 144336: contig of 4011 bp in length
* 144337 144437: gap of unknown length
* 144438 151171: contig of 6735 bp in length
* 151172 151271: gap of unknown length
* 151272 159310: contig of 8039 bp in length
* 159311 159410: gap of unknown length
* 159411 168312: contig of 8902 bp in length
* 168313 168412: gap of unknown length
* 168413 184375: contig of 15963 bp in length.
FEATURES
*
source
1.184375 Location/Qualifiers
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="3"
    /map="3p"
    /clone="RP11-815J19"
misc_feature
1.1389
    /note="assembly_name:Contig35"
misc_feature
1490.3602
    /note="assembly_name:Contig37"
misc_feature
3703.6147
    /note="assembly_name:Contig42"
misc_feature
6248.7841
    /note="assembly_name:Contig45"
misc_feature
7942.9569
    /note="assembly_name:Contig48"
misc_feature
9670.11970
    /note="assembly_name:Contig50"
misc_feature
12071.13209
    /note="assembly_name:Contig51"

```

```

misc_feature 13310.14906
    /note="assembly_name:Contig52"
misc_feature 15007.16915
    /note="assembly_name:Contig53"
misc_feature 17016.18655

```

```

Query Match 1.3%: Score 37; DB 2; Length 184375;
Best local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2462 atgagaaacctgtcccaattgtgtgtcttcctc 2498
|||||
DB 158390 ATTGAGAACCTGTCTCCATTGTGTCTTCCTC 158426

```

RESULT 14

```

AC027579 148270 bp DNA linear HTG 23-JUN-2000
LOCUS Homo sapiens chromosome 17 clone RP11-177D1 map 17, WORKING DRAFT
DEFINITION SEQUENCE, 13 unordered pieces.
ACCESSION AC027579
VERSION AC027579.2 GI:8671976
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

```

REFERENCE 1 (bases 1 to 148270)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
JOURNAL Homo sapiens chromosome 17, clone RP11-177D1
REFERENCE 2 (bases 1 to 148270)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Bouckgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choedel,I., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-pierre,N., Grant,G., Hages,B., Heathord,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Lakoque,K., Lamazares,R., Landers,T., Lenoczky,J.,
Levine,R., Liu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R.,
Meldrim,J., Menues,L., Mihoval,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

```

```

TITLE Direct Submission
JOURNAL Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
RESEARCH, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Jun 23, 2000 this sequence version replaced gi:7342324.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

```

```

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu

```

```

----- Project Information
Center project name: L9001
Center clone name: L77_D_1

```

```

----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 141141 bases at least Q40

```

Consensus quality: 144792 bases at least Q30
Consensus quality: 146242 bases at least Q20
Insert size: 151000; agarose-fp
Insert size: 147070; sum-of-ctnigs
Quality coverage: 4.7 in Q20 bases; agarose-fp
Quality coverage: 4.8 in Q20 bases; sum-of-ctnigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 131: contig of 131 bp in length
132 231: gap of 100 bp
232 4256: contig of 4025 bp in length
4257 4356: gap of 100 bp
4357 6715: contig of 2359 bp in length
6716 6815: gap of 100 bp
6816 12854: contig of 6039 bp in length
12855 12954: gap of 100 bp
12955 18900: contig of 5946 bp in length
18901 19000: gap of 100 bp
19001 27560: contig of 8560 bp in length
27561 27660: gap of 100 bp
27661 41935: contig of 14275 bp in length
41936 42035: gap of 100 bp
42036 58175: contig of 16140 bp in length
58176 58275: gap of 100 bp
58276 70930: contig of 12655 bp in length
70931 71030: gap of 100 bp
71031 87632: contig of 16602 bp in length
87633 87732: gap of 100 bp
87733 104759: contig of 17027 bp in length
104760 104859: gap of 100 bp
104860 123863: contig of 19004 bp in length
123864 123963: gap of 100 bp
123964 148270: contig of 24307 bp in length.

FEATURES

source
1. 148270
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-177D1"
/clone_1lb="RPCT-11 Human Male BAC"
1. 131
/note="assembly_fragment"
clone_end:17
vector_side:right"
232. 4256
/note="assembly_fragment"
4357. 6715
/note="assembly_fragment"
clone_end:SP6
vector_side:right"
6816. 12854
/note="assembly_fragment"
12955. 18900
/note="assembly_fragment"
19001. 27560
/note="assembly_fragment"
27661. 41935
/note="assembly_fragment"
42036. 58175
/note="assembly_fragment"
58276. 70930
/note="assembly_fragment"
71031. 87632
/note="assembly_fragment"
87733. 104759

misc_feature /note="assembly_fragment"
104860. 123863
/note="assembly_fragment"
misc_feature 123964. 148270
/note="assembly_fragment"
BASE COUNT 40333 a 32099 c 32432 g 42206 t 1200 others
ORIGIN

Query Match 138; Score 36; DB 2; Length 148270;
Best Local Similarity 100.0%; Pred. No. 9,1e-09;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2462 attgagaccgtcttcctccattgttgcttctc 2497
|||||
Db 76167 ATTGAGAACCGTCCTTCCCAATTGCTGCTTCTCT 76202

RESULT 15
AC022727
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
human.
HTG: HTGS_PHASE1; HTGS_DRAFT.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
2 (bases 1 to 174405)
Britten,B., Linton,L., Nussbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barne,N., Beckerly,R., Bede,F.,
Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArillano,K., Dewar,K., Domino,M., Doyle,M., Fenebor,J.,
Ferrira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howard,J.C., Johnson,R., Jones,C., Kann,L., Karats,A., Klein,J.,
Lander,T., Lehocsky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marguis,N., McEwan,P., McGurk,A., McEwan,R.,
McPheeters,R., Meldrim,J., Menes,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Tirelli,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.

TITLE
JOURNAL
COMMENT
Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 25, 2000 this sequence version replaced g1:7658384.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5488
Center clone name: 308_J14
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 169089 bases at least Q40
Consensus quality: 172164 bases at least Q30

Consensus quality: 173032 bases at least Q20
 Insert size: 177000; agarose-fp
 Insert size: 173405; sum-of-contigs
 Quality coverage: 4.8 in Q20 bases; agarose-fp
 Quality coverage: 4.9 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 11 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

* 1 6450: contig of 6450 bp in length
* 6451 6550: gap of 100 bp
* 6551 12911: contig of 6361 bp in length
* 12912 13011: gap of 100 bp
* 13012 22057: contig of 9046 bp in length
* 22058 22157: gap of 100 bp
* 22158 32841: contig of 10684 bp in length
* 32842 32941: gap of 100 bp
* 32942 48367: contig of 15426 bp in length
* 48368 48467: gap of 100 bp
* 48468 63919: contig of 15452 bp in length
* 63920 64019: gap of 100 bp
* 64020 78259: contig of 14240 bp in length
* 78260 78359: gap of 100 bp
* 78360 98415: contig of 20056 bp in length
* 98416 98515: gap of 100 bp
* 98516 121876: contig of 23361 bp in length
* 121877 121976: gap of 100 bp
* 121977 144489: contig of 22513 bp in length
* 144490 144589: gap of 100 bp
* 144590 174405: contig of 29816 bp in length.

```

FEATURES
 Source
 1. 174405
 Location/Qualifiers

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-308J14"
1. 6450
/misc_feature
  /note="assembly_fragment"
  6551..12911
/misc_feature
  /note="assembly_fragment"
  13012..22057
/misc_feature
  /note="assembly_fragment"
  22158..32841
/misc_feature
  /note="assembly_fragment"
  32942..48367
/misc_feature
  /note="assembly_fragment"
  48468..63919
/misc_feature
  /note="assembly_fragment"
  64020..78259
/misc_feature
  /note="assembly_fragment"
  clone_end:SP6
  vector_side:right"
  78360..98415
/misc_feature
  /note="assembly_fragment"
  clone_end:T7
  vector_side:left"
  98516..121876
/misc_feature
  /note="assembly_fragment"
  121977..144489
/misc_feature
  /note="assembly_fragment"
  144590..174405
  /note="assembly_fragment"
BASE COUNT 46895 a 38223 c 37328 g 50956 t 1003 others
ORIGIN

```

Query Match 1.3%; Score 36; DB 2; Length 174405;
 Best Local Similarity 100.0%; Pred. No. 9e-09;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2462 attgagacctgtctccattgtgtgtcttct 2497
 Db 127497 ATTGAGACCTGTCTCCATTGTGTGTCTTCT 127532

RESULT 16

AP001531 179792 bp DNA linear HTG 08-AUG-2000
 LOCUS Homo sapiens chromosome 18 clone RP11-772F18 map 18p11.3, WORKING
 DEFINITION DRAFT SEQUENCE, 24 unordered pieces.
 ACCESSION AP001531
 VERSION AP001531.3 GI:9757457
 KEYWORDS HTG, HTGS, PHASE1, HTGS, DRAFT,
 SOURCE Homo sapiens DNA, clone:RP11-772F18.
 ORGANISM Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 179792)

TITLE

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Homo sapiens 179,792 genomic DNA of 18p11.3
 Published Only in Database (2000) In press

JOURNAL

2 (bases 1 to 179792)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Direct Submission
 Submitted (22-MAR-2000) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan
 E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/
 Tel:81-42-778-9923, Fax:81-42-778-9924
 On Aug 9, 2000 this sequence version replaced gi:8117372.

COMMENT

----- Genome Center
 Center: RIKEN Genomic Sciences Center(GSC)
 Center code: RIKEN
 Web site: http://hgp.gsc.riken.go.jp/
 Contact: hattori@gsc.riken.go.jp
 Project Information
 Center project name: RP11-772F18
 Center clone name: RP11-772F18

----- Summary Statistics
 Sequencing vector: PCR products; 100% of reads
 Chemistry: Dye-terminator ET-amersham; 100% of reads Assembly
 Program: Phrap; version 0.990329
 Consensus quality: 171445 bases at least Q40
 Consensus quality: 173267 bases at least Q30
 Consensus quality: 175231 bases at least Q20
 Insert size: 177492; sum-of-contigs
 Quality coverage: 9.04x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
 24 contigs. The true order of the pieces is not known and their
 order in this sequence record is arbitrary. Gaps between the
 contigs are represented as runs N, but the exact sizes of the gaps
 are unknown. This record will be updated with the finished sequence
 as soon as it is available and the accession number will be
 preserved

```

1 24375 contig of 24375 bp in length
24476 45595 contig of 21120 bp in length
45696 55640 contig of 9945 bp in length
55741 70072 contig of 14332 bp in length
70173 80776 contig of 10604 bp in length
80877 89725 contig of 8849 bp in length
89826 99984 contig of 10159 bp in length
100085 108601 contig of 8517 bp in length
108702 117836 contig of 9135 bp in length
117937 125704 contig of 7768 bp in length
125805 134648 contig of 9044 bp in length
134949 142692 contig of 7744 bp in length

```

142793 151024 contig of 8232 bp in length
151125 157334 contig of 6210 bp in length
157435 162993 contig of 5559 bp in length
163094 167067 contig of 3974 bp in length
167168 169460 contig of 2293 bp in length
169561 171662 contig of 2102 bp in length
171763 173150 contig of 1388 bp in length
173251 175114 contig of 1864 bp in length
175215 176405 contig of 1191 bp in length
175506 177563 contig of 1058 bp in length
177664 178673 contig of 1010 bp in length
178774 179792 contig of 1019 bp in length.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 24375: contig of 24375 bp in length
24376 24475: gap of 100 bp
24476 45595: contig of 21120 bp in length
45596 45695: gap of 100 bp
45696 55640: contig of 9945 bp in length
55641 55740: gap of 100 bp
55741 70072: contig of 14332 bp in length
70073 70172: gap of 100 bp
70173 80776: contig of 10604 bp in length
80777 80876: gap of 100 bp
80877 89725: contig of 8849 bp in length
89726 89825: gap of 100 bp
89826 99984: contig of 10159 bp in length
99985 100084: gap of 100 bp
100085 108601: contig of 8517 bp in length
108602 108701: gap of 100 bp
108702 117836: contig of 9135 bp in length
117837 117936: gap of 100 bp
117937 125704: contig of 7768 bp in length
125705 125804: gap of 100 bp
125805 134848: contig of 9044 bp in length
134849 134948: gap of 100 bp
134949 142692: contig of 7744 bp in length
142693 142792: gap of 100 bp
142793 151024: contig of 8232 bp in length
151025 151124: gap of 100 bp
151125 157334: contig of 6210 bp in length
157335 157434: gap of 100 bp
157435 162993: contig of 5559 bp in length
162994 163093: gap of 100 bp
163094 167067: contig of 3974 bp in length
167068 167167: gap of 100 bp
167168 169460: contig of 2293 bp in length
169461 169560: gap of 100 bp
169561 171662: contig of 2102 bp in length
171663 171762: gap of 100 bp
171763 173150: contig of 1388 bp in length
173151 173250: gap of 100 bp
173251 175114: contig of 1864 bp in length
175115 175214: gap of 100 bp
175215 176405: contig of 1191 bp in length
176406 176505: gap of 100 bp
176506 177563: contig of 1058 bp in length
177564 177663: gap of 100 bp
177664 178673: contig of 1010 bp in length
178674 178773: gap of 100 bp
178774 179792: contig of 1019 bp in length.

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18p11.3"

/clone="RP11-772F18"
1. 24375
/note="assembly-fragment"
24476..45595
/note="assembly-fragment"
45696..55640
/note="assembly-fragment"
55741..70072
/note="assembly-fragment"
70173..80776
/note="assembly-fragment"
80877..89725
/note="assembly-fragment"
89826..99984
/note="assembly-fragment"
100085..108601
/note="assembly-fragment"
108702..117836
/note="assembly-fragment"
117937..125704
/note="assembly-fragment"
125805..134848
/note="assembly-fragment"
134949..142692
/note="assembly-fragment"
142793..151024
/note="assembly-fragment"
151125..157334
/note="assembly-fragment clone_end:SP6 vector_side:right"
157435..162993
/note="assembly-fragment"
163094..167067
/note="assembly-fragment"
167168..169460
/note="assembly-fragment clone_end:T7 vector_side:left"
169561..171662
/note="assembly-fragment"
171763..173150
/note="assembly-fragment"
173251..175114
/note="assembly-fragment"
175215..176405
/note="assembly-fragment"
176506..177563
/note="assembly-fragment"
177664..178673
/note="assembly-fragment"
178774..179792
/note="assembly-fragment"
BASE COUNT 49430 a 37736 c 39365 g 50954 t 2307 others
ORIGIN

Query Match 1.3%: Score 36; DB 2; Length 179792;
Best Local Similarity 100.0%; Pred. No. 9e-09;

Matches .36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2462 attgaagacctgtctccattgtgtcttcct 2497
|||||
Db 29905 ATTGAGACCTGTCTCCATTGTGTCTTCTT 29940

RESULT 17

AC025211

LOCUS AC025211 150794 bp DNA linear PRI 08-SEP-2001

DEFINITION Homo sapiens chromosome , clone RP11-220C2, complete sequence.

ACCESSION AC025211

VERSION AC025211.8 GI:15528921

KEYWORDS

SOURCE

ORGANISM

Homo sapiens
human.
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.


```

repeat_region      /rpt_family="L2"
                    complement(17671..17984)
                    /rpt_family="AluJb"
repeat_region      complement(18115..18681)
                    /rpt_family="L1ME1"
repeat_region      complement(18783..19023)
                    /rpt_family="L1ME1"
repeat_region      complement(19024..19321)
                    /rpt_family="AluSg"
repeat_region      complement(19322..20092)
                    /rpt_family="L1ME1"
repeat_region      20098..20194
                    /rpt_family="Cr-rich"
repeat_region      complement(20195..20356)
                    /rpt_family="AluJb"
repeat_region      20357..20395
                    /rpt_family="(TGT)n"
repeat_region      complement(20396..20690)
                    /rpt_family="AluSx"
repeat_region      complement(20691..20830)
                    /rpt_family="AluJb"
repeat_region      complement(20831..22342)
                    /rpt_family="L1ME1"
repeat_region      complement(21587..21602)
                    /note="<30 qual SNGU region"
repeat_region      complement(21644..21658)
                    /note="<30 qual SNGU region"
repeat_region      complement(22343..22535)
                    /rpt_family="AluJb"
repeat_region      complement(22629..22718)
                    /rpt_family="AluJ/FLAM"
repeat_region      complement(23036..23116)
                    /rpt_family="L1MB2"
repeat_region      complement(23117..23399)
                    /rpt_family="AluJo"
repeat_region      23294..23299
                    /note="<30 qual SNGU region"
repeat_region      complement(23400..23422)
                    /rpt_family="L1MB2"
repeat_region      complement(23472..23561)
                    /rpt_family="MIR"
repeat_region      24063..24374
                    /rpt_family="AluSx"
repeat_region      complement(24989..25192)
                    /rpt_family="L3"

Query Match      1.2%: Score 34: DB 9: Length 150794:
Best Local Similarity 100.0%: Pred No. 1.3e-07:
Matches 34: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 2464 tgaagacctgtctccatttgatgcttctct 2497
      |||
Db 109211 TGAGAACCTGCTCCATTGTGCTCTTCC 109244

RESULT 18
AC087493      166650 bp      DNA      linear      HTG 24-MAY-2001
LOCUS      Homo sapiens chromosome 17 clone RP11-253N19 map 17, WORKING DRAFT
DEFINITION
AC087493      AC087493.3      GI:14192963
VERSION      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
SOURCE      Homo sapiens
              human.
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              1 (bases 1 to 166650)
REFERENCE      Birren,B., Linton,L., Nusbaum,C. and Lander,E.
              TITLE      Homo sapiens chromosome 17, clone RP11-253N19
              JOURNAL      Unpublished
              2 (bases 1 to 166650)
              Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,

```

```

TITLE
JOURNAL
COMMENT
COMMENT
Submitted (05-JAN-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 24, 2001 this sequence version replaced gi:1312325.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L11973
Center clone name: 253_N_19

----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 165769 bases at least Q40
Consensus quality: 166122 bases at least Q30
Consensus quality: 166247 bases at least Q20
Insert size: 163000; agarose-fp
Insert size: 166350; sum-of-ctrls
Quality coverage: 11.0 in Q20 bases; agarose-fp
Quality coverage: 10.8 in Q20 bases; sum-of-ctrls

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 11793: contig of 11793 bp in length
*
* 11794 11893: gap of 100 bp
*
* 11894 33667: contig of 21774 bp in length
*
* 33668 33767: gap of 100 bp
*
* 33768 97497: contig of 63730 bp in length
*
* 97498 97597: gap of 100 bp
*
* 97598 166650: contig of 69053 bp in length.
*
Location/Qualifiers
1..166650
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-253N19"
/clone_1lb="RPC1-11 Human Male BAC"
1..11793
/note="assembly-fragment
clone_end:SP6

```

```

misc-feature      vector_side:left"
                  11894..33667
                  /note="assembly_fragment"
misc-feature      33768..97497
                  /note="assembly_fragment"
misc-feature      97598..166650
                  /note="assembly_fragment"
                  clone_end:17
                  vector_side:right"

```

```

BASE COUNT      51731 a 33076 c 32370 g 49168 t 305 others
ORIGIN

```

```

Query Match      1.2%; Score 34; DB 2; Length 166650;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2464 ttgagaacctgtctccattgtgtgtcttctcct 2497
      |||
Db 39715 ttgagaacctgtctccattgtgtgtcttctcct 39682

```

```

RESULT 19
AC008106      218922 bp      DNA      linear      HTG 04-SEP-2000
LOCUS      Homo sapiens chromosome 17 clone CTD-3008D3 map 17, WORKING DRAFT
DEFINITION
ACCESSION      AC008106
VERSION      AC008106.3 GI:9966323
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM      Homo sapiens

```

```

REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE      1 (bases 1 to 218922)
JOURNAL      Homo sapiens chromosome 17, clone CTD-3008D3
REFERENCE      Unpublished
AUTHORS      2 (bases 1 to 218922)

```

```

Baird, J., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., DeFelicio, K., Depayre, E., Devon, K., Dewar, K.,
Donelan, L., Doyle, M., Ferreira, P., FitzHugh, M., Forrest, C.,
Funke, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G.,
Hagos, B., Headford, A., Horton, L., Howland, J. C., Jones, C., Kann, L.,
Karatas, A., Lehoczeky, J., Liu, C., Locke, K., Macdonald, P.,
Margulis, N., McEwan, P., McGuire, A., McKernan, K., McLaughlin, J.,
Meldrum, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,
Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B.,
Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A.,
Testaye, S., Toruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A.,
Wheeler, J., Wu, X., Wyman, D., Ye, W. J., and Zody, M.

```

```

COMMENT
TITLE      Direct Submission
JOURNAL      Submitted (22-JUL-1999) Whitehead Institute/MIT Center for Genome
              Research, 320 Charles Street, Cambridge, MA 02141, USA
              On Sep 4, 2000 this sequence version replaced gi:7329486.
              All repeats were identified using RepeatMasker:
              Smit, A.F.A. & Green, P. (1996-1997)
              http://ftp.genome.washington.edu/RM/RepeatMasker.html

```

```

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR

```

```

Web site: http://www-seq.wi.mit.edu

```

```

Contact: sequence_submissions@genome.wi.mit.edu

```

```

----- Project Information

```

```

Center project name: L900

```

```

Center clone name: 3008_D_3

```

```

----- Summary Statistics

```

```

Sequencing vector: M13; M77815; 100% of reads

```

```

Chemistry: Dye primer; amersham; 100% of reads

```

```

Assembly program: Phrap; version 0.960731

```

```

Consensus quality: 197177 bases at least Q40
Consensus quality: 206822 bases at least Q30
Consensus quality: 211643 bases at least Q20
Insert size: 200000; agarose-fp
Insert size: 215622; sum-of-contigs
Quality coverage: 4.1 in Q20 bases; agarose-fp
Quality coverage: 3.8 in Q20 bases; sum-of-contigs

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

1 1091 1190: contig of 1090 bp in length
* 1091 1190: gap of 100 bp
* 1191 2690: contig of 1500 bp in length
* 2691 2790: gap of 100 bp
* 2791 4012: contig of 1222 bp in length
* 4013 4112: gap of 100 bp
* 4113 5123: contig of 1011 bp in length
* 5124 5223: gap of 100 bp
* 5224 6312: contig of 1089 bp in length
* 6313 6412: gap of 100 bp
* 6413 8474: contig of 2062 bp in length
* 8475 8574: gap of 100 bp
* 8575 10328: contig of 1753 bp in length
* 10329 10428: gap of 100 bp
* 10429 12116: contig of 1688 bp in length
* 12117 12216: gap of 100 bp
* 12217 14183: contig of 1967 bp in length
* 14184 14283: gap of 100 bp
* 14284 17330: contig of 3047 bp in length
* 17331 17430: gap of 100 bp
* 17431 21134: contig of 3704 bp in length
* 21135 21234: gap of 100 bp
* 21235 24428: contig of 3194 bp in length
* 24429 24528: gap of 100 bp
* 24529 27705: contig of 3177 bp in length
* 27706 27805: gap of 100 bp
* 27806 30106: contig of 2301 bp in length
* 30107 30206: gap of 100 bp
* 30207 35366: contig of 5160 bp in length
* 35367 35466: gap of 100 bp
* 35467 39634: contig of 4168 bp in length
* 39635 39734: gap of 100 bp
* 39735 43767: contig of 4033 bp in length
* 43768 43867: gap of 100 bp
* 43868 50104: contig of 6237 bp in length
* 50105 50204: gap of 100 bp
* 50205 54896: contig of 4692 bp in length
* 54897 54996: gap of 100 bp
* 54997 61677: contig of 6681 bp in length
* 61678 61777: gap of 100 bp
* 61778 68042: contig of 6265 bp in length
* 68043 68142: gap of 100 bp
* 68143 76003: contig of 7861 bp in length
* 76004 76103: gap of 100 bp
* 76104 83935: contig of 7832 bp in length
* 83936 84035: gap of 100 bp
* 84036 92731: contig of 8696 bp in length
* 92732 92831: gap of 100 bp
* 92832 102205: contig of 9374 bp in length
* 102206 102305: gap of 100 bp
* 102306 111555: contig of 9250 bp in length
* 111556 111655: gap of 100 bp
* 111656 120772: contig of 9117 bp in length
* 120773 120872: gap of 100 bp
* 120873 129869: contig of 8897 bp in length
* 129870 129969: gap of 100 bp
* 129970 143160: contig of 13191 bp in length

```

*	143161	143360:	gap of	100 bp
*	143361	154541:	contig of	11281 bp
*	154361	154541:	contig of	100 bp
*	154362	154541:	gap of	100 bp
*	154442	154541:	gap of	100 bp
*	154443	154541:	gap of	100 bp
*	154444	154541:	gap of	100 bp
*	154445	154541:	gap of	100 bp
*	154446	154541:	gap of	100 bp
*	154447	154541:	gap of	100 bp
*	154448	154541:	gap of	100 bp
*	154449	154541:	gap of	100 bp
*	154450	154541:	gap of	100 bp
*	154451	154541:	gap of	100 bp
*	154452	154541:	gap of	100 bp
*	154453	154541:	gap of	100 bp
*	154454	154541:	gap of	100 bp
*	154455	154541:	gap of	100 bp
*	154456	154541:	gap of	100 bp
*	154457	154541:	gap of	100 bp
*	154458	154541:	gap of	100 bp
*	154459	154541:	gap of	100 bp
*	154460	154541:	gap of	100 bp
*	154461	154541:	gap of	100 bp
*	154462	154541:	gap of	100 bp
*	154463	154541:	gap of	100 bp
*	154464	154541:	gap of	100 bp
*	154465	154541:	gap of	100 bp
*	154466	154541:	gap of	100 bp
*	154467	154541:	gap of	100 bp
*	154468	154541:	gap of	100 bp
*	154469	154541:	gap of	100 bp
*	154470	154541:	gap of	100 bp
*	154471	154541:	gap of	100 bp
*	154472	154541:	gap of	100 bp
*	154473	154541:	gap of	100 bp
*	154474	154541:	gap of	100 bp
*	154475	154541:	gap of	100 bp
*	154476	154541:	gap of	100 bp
*	154477	154541:	gap of	100 bp
*	154478	154541:	gap of	100 bp
*	154479	154541:	gap of	100 bp
*	154480	154541:	gap of	100 bp
*	154481	154541:	gap of	100 bp
*	154482	154541:	gap of	100 bp
*	154483	154541:	gap of	100 bp
*	154484	154541:	gap of	100 bp
*	154485	154541:	gap of	100 bp
*	154486	154541:	gap of	100 bp
*	154487	154541:	gap of	100 bp
*	154488	154541:	gap of	100 bp
*	154489	154541:	gap of	100 bp
*	154490	154541:	gap of	100 bp
*	154491	154541:	gap of	100 bp
*	154492	154541:	gap of	100 bp
*	154493	154541:	gap of	100 bp
*	154494	154541:	gap of	100 bp
*	154495	154541:	gap of	100 bp
*	154496	154541:	gap of	100 bp
*	154497	154541:	gap of	100 bp
*	154498	154541:	gap of	100 bp
*	154499	154541:	gap of	100 bp
*	154500	154541:	gap of	100 bp
*	154501	154541:	gap of	100 bp
*	154502	154541:	gap of	100 bp
*	154503	154541:	gap of	100 bp
*	154504	154541:	gap of	100 bp
*	154505	154541:	gap of	100 bp
*	154506	154541:	gap of	100 bp
*	154507	154541:	gap of	100 bp
*	154508	154541:	gap of	100 bp
*	154509	154541:	gap of	100 bp
*	154510	154541:	gap of	100 bp
*	154511	154541:	gap of	100 bp
*	154512	154541:	gap of	100 bp
*	154513	154541:		

source	FEATURES
--------	----------

misc_feature	/organism="Homo sapiens"
misc_feature	/db_xref="taxon:9606"
misc_feature	/chromosome="17"
misc_feature	/map="17"
misc_feature	/clone="CTD-3008D3"
misc_feature	/clone_lib="CTD Human BAC"
misc_feature	1..1090
misc_feature	/note="assembly_fragment"
misc_feature	1191..2690
misc_feature	/note="assembly_fragment"
misc_feature	2791..4012
misc_feature	/note="assembly_fragment"
misc_feature	4113..5123
misc_feature	/note="assembly_fragment"
misc_feature	5224..6312
misc_feature	/note="assembly_fragment"
misc_feature	6413..8474
misc_feature	/note="assembly_fragment"
misc_feature	8575..10328
misc_feature	/note="assembly_fragment"
misc_feature	10429..12116
misc_feature	/note="assembly_fragment"
misc_feature	12217..14183
misc_feature	/note="assembly_fragment"
misc_feature	14284..17330
misc_feature	/note="assembly_fragment"
misc_feature	17431..21134
misc_feature	/note="assembly_fragment"
misc_feature	21235..24428
misc_feature	/note="assembly_fragment"
misc_feature	24529..27705
misc_feature	/note="assembly_fragment"
misc_feature	27806..30106
misc_feature	/note="assembly_fragment"
misc_feature	30207..35366
misc_feature	/note="assembly_fragment"
misc_feature	35467..39634
misc_feature	/note="assembly_fragment"
misc_feature	39735..43767
misc_feature	/note="assembly_fragment"
misc_feature	43868..50104
misc_feature	/note="assembly_fragment"
misc_feature	50205..54896
misc_feature	/note="assembly_fragment"
misc_feature	54997..61677
misc_feature	/note="assembly_fragment"
misc_feature	61778..68042
misc_feature	/note="assembly_fragment"
misc_feature	68143..76003
misc_feature	/note="assembly_fragment"
misc_feature	76104..83935
misc_feature	/note="assembly_fragment"
misc_feature	84036..92731
misc_feature	/note="assembly_fragment"
misc_feature	92832..102205
misc_feature	/note="assembly_fragment"
misc_feature	102306..111555
misc_feature	/note="assembly_fragment"
misc_feature	111656..120772
misc_feature	/note="assembly_fragment"
misc_feature	120873..129869

```
misc_feature /note="assembly_fragment"
129970. 143160
misc_feature /note="assembly_fragment"
143261. 154541
```

Query Match	1.28;	Score 34;	DB 2;	Length 218922;
Best Local Similarity	100.0%;	Pred. No. 1.3e-07;		
Matches 34;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	2464	tgagaacctgtcttccatttgatgtccttcct	2497
Db	99714	TGAGACCTGTCTTCCATTGGTGTCCTTCTCCT	99747

RESULT 20
AC027607/0

LOCUS	162453 bp	DNA	linear	HTG 31-AUG-2001
DEFINITION	Homo sapiens chromosome 4 clone RP11-368B15, WORKING			
DESCRIPTION	3 unordered pieces.			

ACCESSION	AC027607
VERSION	AC027607.6
KEYWORDS	GI:15321552
SOURCE	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN human.

REFERENCE

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 162453)

REFERENCE

The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 162453)

JOURNAL

Direct Submission
Submitted (30-MAR-2000) Genome Sequencing Center, Washington

COMMENT

University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Aug 28, 2001 this sequence version replaced gi:15145259.

```
----- genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0036815
```

Summary Statistics

Sequencing vector: M13: 79%
Sequencing vector: plasmid: 21%
Chemistry: Dye-Primer ET: 79% of reads
Chemistry: Dye-terminator Big Dye: 21% of reads
Assembly program: Phrap: version 0.990319
Consensus quality: 162283 bases at least Q40
Consensus quality: 106660 bases at least Q30
Consensus quality: 168288 bases at least Q20
Insert size: 146000; agarose-1p

Quality coverage: 9.33 in Q20 bases; agarose-fp
Quality coverage: 8.80 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

*	1	1192: contig of 1192 bp in length
*	1193	1292: gap of unknown length
*	1293	72574: contig of 71282 bp in length
*	72575	72674: gap of unknown length
*	72675	152455: contig of 89779 bp in length

REFERENCE
AUTHORS

2 (bases 1 to 173133)
Baird, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Balwin, J., Barna, N., Beckley, R., Boguslavsky, L., Boukhalter, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Coore, P., Deatellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M., Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Galand, J., Gardyna, S., Grant, G., Hagos, B., Hearford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lechoczky, J., Lien, C., Locke, K., MacDonald, P., Margolis, N., McEwan, P., McGuirk, A., McKernan, K., McLaughlin, T., Meltrin, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testafaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

TITLE
JOURNAL

COMMENT

Direct Submission
Submitted (23-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 20, 2000 this sequence version replaced gi:9108714.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu

Project Information
Center project name: L1434
Center clone name: 1_D_6

Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 156896 bases at least Q40
Consensus quality: 165527 bases at least Q40
Consensus quality: 168595 bases at least Q20
Insert size: 194000; agarose-fp
Insert size: 170133; sum-of-contigs
Quality coverage: 3.5 in Q20 bases; agarose-fp
Quality coverage: 4.0 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 31 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 5114: contig of 5114 bp in length
* 5115 5214: gap of 100 bp
* 5215 6404: contig of 1190 bp in length
* 6405 6504: gap of 100 bp
* 6505 7651: contig of 1147 bp in length
* 7652 7751: gap of 100 bp
* 7752 9355: contig of 1604 bp in length
* 9356 9455: gap of 100 bp
* 9456 10655: contig of 1200 bp in length
* 10656 10755: gap of 100 bp
* 10756 11937: contig of 1182 bp in length
* 11938 12037: gap of 100 bp
* 12038 13055: contig of 1018 bp in length
* 13056 13155: gap of 100 bp
* 13156 14204: contig of 1049 bp in length
* 14205 14304: gap of 100 bp
* 14305 15800: contig of 1496 bp in length
* 15801 15900: gap of 100 bp
* 15901 17319: contig of 1419 bp in length
* 17320 17419: gap of 100 bp
* 17420 19622: contig of 2203 bp in length
* 19623 19722: gap of 100 bp
* 19723 21692: contig of 1970 bp in length

FEATURES

source

misc_feature

21693 21792: gap of 100 bp
* 21793 23722: contig of 1930 bp in length
* 23723 23822: gap of 100 bp
* 23823 27350: contig of 3528 bp in length
* 27351 27450: gap of 100 bp
* 27451 30128: contig of 2678 bp in length
* 30129 30228: gap of 100 bp
* 30229 33435: contig of 3207 bp in length
* 33436 33535: gap of 100 bp
* 33536 36762: contig of 3227 bp in length
* 36763 36862: gap of 100 bp
* 36863 41461: contig of 4599 bp in length
* 41462 41561: gap of 100 bp
* 41562 45986: contig of 4425 bp in length
* 45987 46086: gap of 100 bp
* 46087 51166: contig of 5080 bp in length
* 51167 51266: gap of 100 bp
* 51267 55890: contig of 4624 bp in length
* 55891 55990: gap of 100 bp
* 55991 62876: contig of 6886 bp in length
* 62877 62976: gap of 100 bp
* 62977 70430: contig of 7454 bp in length
* 70431 70530: gap of 100 bp
* 70531 87882: contig of 17352 bp in length
* 87883 87982: gap of 100 bp
* 87983 97567: contig of 9585 bp in length
* 97568 97667: gap of 100 bp
* 97668 106038: contig of 8371 bp in length
* 106039 106138: gap of 100 bp
* 106139 118960: contig of 12822 bp in length
* 118961 119060: gap of 100 bp
* 119061 133793: contig of 14733 bp in length
* 133794 133893: gap of 100 bp
* 133894 148302: contig of 14809 bp in length
* 148303 148402: gap of 100 bp
* 148403 165690: contig of 17288 bp in length
* 165691 165900: gap of 100 bp
* 16591 173133: contig of 7343 bp in length.
* 165791

Location/Qualifiers

1. 173133
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-106"
/clone_lib="RP11 Human Male BAC"
1. 5114
/note="assembly_fragment"
clone_end:SP6
vector_side:left
5215..6404
/note="assembly_fragment"
6505..7651
/note="assembly_fragment"
7752..9355
/note="assembly_fragment"
9456..10655
/note="assembly_fragment"
10756..11937
/note="assembly_fragment"
12038..13055
/note="assembly_fragment"
13156..14204
/note="assembly_fragment"
14305..15800
/note="assembly_fragment"
15901..17319
/note="assembly_fragment"
17420..19622
/note="assembly_fragment"
19723..21692
/note="assembly_fragment"
21793..23722
/note="assembly_fragment"
23823..27350

misc_feature

21693 21792: gap of 100 bp
* 21793 23722: contig of 1930 bp in length
* 23723 23822: gap of 100 bp
* 23823 27350: contig of 3528 bp in length
* 27351 27450: gap of 100 bp
* 27451 30128: contig of 2678 bp in length
* 30129 30228: gap of 100 bp
* 30229 33435: contig of 3207 bp in length
* 33436 33535: gap of 100 bp
* 33536 36762: contig of 3227 bp in length
* 36763 36862: gap of 100 bp
* 36863 41461: contig of 4599 bp in length
* 41462 41561: gap of 100 bp
* 41562 45986: contig of 4425 bp in length
* 45987 46086: gap of 100 bp
* 46087 51166: contig of 5080 bp in length
* 51167 51266: gap of 100 bp
* 51267 55890: contig of 4624 bp in length
* 55891 55990: gap of 100 bp
* 55991 62876: contig of 6886 bp in length
* 62877 62976: gap of 100 bp
* 62977 70430: contig of 7454 bp in length
* 70431 70530: gap of 100 bp
* 70531 87882: contig of 17352 bp in length
* 87883 87982: gap of 100 bp
* 87983 97567: contig of 9585 bp in length
* 97568 97667: gap of 100 bp
* 97668 106038: contig of 8371 bp in length
* 106039 106138: gap of 100 bp
* 106139 118960: contig of 12822 bp in length
* 118961 119060: gap of 100 bp
* 119061 133793: contig of 14733 bp in length
* 133794 133893: gap of 100 bp
* 133894 148302: contig of 14809 bp in length
* 148303 148402: gap of 100 bp
* 148403 165690: contig of 17288 bp in length
* 165691 165900: gap of 100 bp
* 16591 173133: contig of 7343 bp in length.
* 165791

Location/Qualifiers

1. 173133
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-106"
/clone_lib="RP11 Human Male BAC"
1. 5114
/note="assembly_fragment"
clone_end:SP6
vector_side:left
5215..6404
/note="assembly_fragment"
6505..7651
/note="assembly_fragment"
7752..9355
/note="assembly_fragment"
9456..10655
/note="assembly_fragment"
10756..11937
/note="assembly_fragment"
12038..13055
/note="assembly_fragment"
13156..14204
/note="assembly_fragment"
14305..15800
/note="assembly_fragment"
15901..17319
/note="assembly_fragment"
17420..19622
/note="assembly_fragment"
19723..21692
/note="assembly_fragment"
21793..23722
/note="assembly_fragment"
23823..27350

misc_feature

21693 21792: gap of 100 bp
* 21793 23722: contig of 1930 bp in length
* 23723 23822: gap of 100 bp
* 23823 27350: contig of 3528 bp in length
* 27351 27450: gap of 100 bp
* 27451 30128: contig of 2678 bp in length
* 30129 30228: gap of 100 bp
* 30229 33435: contig of 3207 bp in length
* 33436 33535: gap of 100 bp
* 33536 36762: contig of 3227 bp in length
* 36763 36862: gap of 100 bp
* 36863 41461: contig of 4599 bp in length
* 41462 41561: gap of 100 bp
* 41562 45986: contig of 4425 bp in length
* 45987 46086: gap of 100 bp
* 46087 51166: contig of 5080 bp in length
* 51167 51266: gap of 100 bp
* 51267 55890: contig of 4624 bp in length
* 55891 55990: gap of 100 bp
* 55991 62876: contig of 6886 bp in length
* 62877 62976: gap of 100 bp
* 62977 70430: contig of 7454 bp in length
* 70431 70530: gap of 100 bp
* 70531 87882: contig of 17352 bp in length
* 87883 87982: gap of 100 bp
* 87983 97567: contig of 9585 bp in length
* 97568 97667: gap of 100 bp
* 97668 106038: contig of 8371 bp in length
* 106039 106138: gap of 100 bp
* 106139 118960: contig of 12822 bp in length
* 118961 119060: gap of 100 bp
* 119061 133793: contig of 14733 bp in length
* 133794 133893: gap of 100 bp
* 133894 148302: contig of 14809 bp in length
* 148303 148402: gap of 100 bp
* 148403 165690: contig of 17288 bp in length
* 165691 165900: gap of 100 bp
* 16591 173133: contig of 7343 bp in length.
* 165791

Location/Qualifiers

1. 173133
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-106"
/clone_lib="RP11 Human Male BAC"
1. 5114
/note="assembly_fragment"
clone_end:SP6
vector_side:left
5215..6404
/note="assembly_fragment"
6505..7651
/note="assembly_fragment"
7752..9355
/note="assembly_fragment"
9456..10655
/note="assembly_fragment"
10756..11937
/note="assembly_fragment"
12038..13055
/note="assembly_fragment"
13156..14204
/note="assembly_fragment"
14305..15800
/note="assembly_fragment"
15901..17319
/note="assembly_fragment"
17420..19622
/note="assembly_fragment"
19723..21692
/note="assembly_fragment"
21793..23722
/note="assembly_fragment"
23823..27350

```

misc_feature      /note="assembly_fragment"
27451..30128
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
30229..33435
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
33536..36762
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
36863..41461
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
41562..45986
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
46087..51166
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
51267..55890
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
55991..62876
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
62977..70430
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
70531..87882
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
87983..97567
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
97668..106038
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
106139..118960
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
119061..133793
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
133894..148302
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
148403..165690
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
165791..173133
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
clone_end:T7
vector_side:right"
BASE COUNT      51650 a 33587 c 33772 g 51116 t 3008 others
ORIGIN

```

```

Query Match      1.2%; Score 32; DB 2; Length 173133;
Best Local Similarity 100.0%; Pred. No. 2e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2457 actgaattggaacctgtctccattggtg 2488
|||||
Db 15324 ACTGAATTGAGACCTGTCTCCATTGGTG 15293

```

```

RESULT 24
AC098581      174701 bp      DNA      linear      HTG 25-OCT-2001
LOCUS      Homo sapiens chromosome 2 clone RP11-1D6, WORKING DRAFT SEQUENCE, 9
DEFINITION      unordered pieces.
ACCESSION      AC098581
VERSION      AC098581.1 GI:16418216
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE      1 (bases 1 to 174701)
AUTHORS      Waterston,R.H.
TITLE      The sequence of Homo sapiens clone
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 174701)
AUTHORS      Waterston,R.H.
TITLE      Direct Submission
JOURNAL      Submitted (25-OCT-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT

```

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: H_NH0001D06
----- Summary Statistics -----
Sequencing vector: MJ3: 0%
Sequencing vector: plasmid: 100%
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 171641 bases at least Q40
Consensus quality: 172606 bases at least Q30
Consensus quality: 173237 bases at least Q20
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1406: contig of 1406 bp in length
* 1407 1506: gap of unknown length
* 1507 6809: contig of 5303 bp in length
* 6810 6909: gap of unknown length
* 6910 20466: contig of 13557 bp in length
* 20467 20566: gap of unknown length
* 20567 35544: contig of 14978 bp in length
* 35545 35644: gap of unknown length
* 35645 52465: contig of 16821 bp in length
* 52466 52565: gap of unknown length
* 52566 73259: contig of 20654 bp in length
* 73260 73359: gap of unknown length
* 73360 105109: contig of 31750 bp in length
* 105110 105209: gap of unknown length
* 105210 140153: contig of 34944 bp in length
* 140154 140253: gap of unknown length
* 140254 174701: contig of 34448 bp in length.
Location/Qualifiers
1. 174701
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-1D6"
1. 1406
/note="assembly_name:Contig5"
1507..6809
/note="assembly_name:Contig6"
6910..20466
/note="assembly_name:Contig7"
20567..35544
/note="assembly_name:Contig8"
35645..52465
/note="assembly_name:Contig9
clone_end:T7
vector_side:right"
52566..73259
/note="assembly_name:Contig10"
73360..105109
/note="assembly_name:Contig11"
105210..140153
/note="assembly_name:Contig12"
140254..174701
/note="assembly_name:Contig13"
BASE COUNT      52648 a 34002 c 34946 g 52304 t 801 others
ORIGIN

```

```

FEATURES
source
misc_feature      /note="assembly_name:Contig5"
1507..6809
/note="assembly_name:Contig6"
6910..20466
/note="assembly_name:Contig7"
20567..35544
/note="assembly_name:Contig8"
35645..52465
/note="assembly_name:Contig9
clone_end:T7
vector_side:right"
52566..73259
/note="assembly_name:Contig10"
73360..105109
/note="assembly_name:Contig11"
105210..140153
/note="assembly_name:Contig12"
140254..174701
/note="assembly_name:Contig13"
BASE COUNT      52648 a 34002 c 34946 g 52304 t 801 others
ORIGIN

```


Query Match 1.2%; Score 32; DB 2; Length 174701;
 Best Local Similarity 100.0%; Pred. No. 2e-06;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2457 actgaatgaagaacctgtctccattgtgtg 2488
 |||||||||||||||||||||||||||||
 DB 116104 ACTGAATGAGAACGCTCTCCATTGTGTG 116135

RESULT 25
 AC073412 175833 bp DNA linear HTG 17-AUG-2000
 LOCUS Homo sapiens chromosome 11 clone RP11-56C8, WORKING DRAFT SEQUENCE,
 23 unordered pieces.
 AC073412
 AC073412.3 GI:9838047
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 175833)
 Waterston, R.H.
 TITLE The sequence of Homo sapiens clone
 2 (bases 1 to 175833)
 Waterston, R.H.
 REFERENCE Submitted (16-JUN-2000) Genome Sequencing Center, Washington
 UNIVERSITY School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 COMMENT On Aug 17, 2000 this sequence version replaced g1:9797840.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 ----- Project Information -----
 Center project name: H_NH0056C08
 ----- Summary Statistics -----
 Sequencing vector: MJ3: 100%
 Sequencing vector: plasmid: 0%
 Chemistry: Dye-terminator Big Dye; 0% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 159440 bases at least Q40
 Consensus quality: 165805 bases at least Q30
 Consensus quality: 169022 bases at least Q20
 Insert size: 169000; agarose-fp
 Insert size: 174679; sum-of-contigs
 Quality coverage: 3.64 in Q20 bases; WORKING DRAFT SEQUENCE
 Quality coverage: 3.59 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 23 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1. 1393: contig of 1393 bp in length
 * 1394 1493: gap of unknown length
 * 1494 2618: contig of 1125 bp in length
 * 2619 2718: gap of unknown length
 * 2719 5227: contig of 2509 bp in length
 * 5228 5327: gap of unknown length
 * 5328 7481: contig of 2154 bp in length
 * 7482 7581: gap of unknown length
 * 7582 10623: contig of 3042 bp in length
 * 10624 10723: gap of unknown length
 * 10724 15116: contig of 4393 bp in length
 * 15117 15216: gap of unknown length

15217 18771: contig of 3555 bp in length
 * 18772 18871: gap of unknown length
 * 18872 24865: contig of 5994 bp in length
 * 24866 24965: gap of unknown length
 * 24966 28388: contig of 3423 bp in length
 * 28389 28488: gap of unknown length
 * 28489 32855: contig of 4367 bp in length
 * 32856 32955: gap of unknown length
 * 32956 37473: contig of 4518 bp in length
 * 37474 37573: gap of unknown length
 * 37574 43030: contig of 5457 bp in length
 * 43031 43130: gap of unknown length
 * 43131 49891: contig of 6761 bp in length
 * 49892 49991: gap of unknown length
 * 49992 57329: contig of 7338 bp in length
 * 57330 57429: gap of unknown length
 * 57430 67593: contig of 9330 bp in length
 * 67594 66859: gap of unknown length
 * 66860 75716: contig of 8857 bp in length
 * 75717 75816: gap of unknown length
 * 75817 84841: contig of 9025 bp in length
 * 84842 84941: gap of unknown length
 * 84942 95395: contig of 10454 bp in length
 * 95396 95495: gap of unknown length
 * 95496 110307: contig of 14811 bp in length
 * 110308 110406: gap of unknown length
 * 110407 123845: contig of 13439 bp in length
 * 123846 123945: gap of unknown length
 * 123946 142389: contig of 18454 bp in length
 * 142400 142489: gap of unknown length
 * 142490 157956: contig of 15457 bp in length
 * 157957 158056: gap of unknown length
 * 158057 175833: contig of 17777 bp in length.

FEATURES
 source
 1..175833
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /clone="RP11-56C8"
 1..1393
 /note="assembly_name:Contig15"
 misc_feature
 1494..2618
 /note="assembly_name:Contig18"
 misc_feature
 2719..5227
 /note="assembly_name:Contig20"
 misc_feature
 5328..7481
 /note="assembly_name:Contig21"
 misc_feature
 7582..10623
 /note="assembly_name:Contig22"
 misc_feature
 10724..15116
 /note="assembly_name:Contig23"
 misc_feature
 15217..18771
 /note="assembly_name:Contig24"
 misc_feature
 18872..24865
 /note="assembly_name:Contig25"
 misc_feature
 24966..28388
 /note="assembly_name:Contig26
 clone_end:T7
 vector_side:right"
 28489..32855
 /note="assembly_name:Contig27"
 misc_feature
 32956..37473
 /note="assembly_name:Contig28"
 misc_feature
 37574..43030
 /note="assembly_name:Contig29"
 misc_feature
 43131..49891
 /note="assembly_name:Contig30"
 misc_feature
 49992..57329
 /note="assembly_name:Contig31"
 misc_feature
 57430..66759
 /note="assembly_name:Contig32"
 misc_feature
 66860..75716
 /note="assembly_name:Contig33"

```

misc_feature      75817..84841
                  /note="assembly_name:Contig34"
misc_feature      84942..95395
                  /note="assembly_name:Contig35"
misc_feature      95496..110306
                  /note="assembly_name:Contig36"
misc_feature      110407..123845
                  /note="assembly_name:Contig37"
misc_feature      123946..142399
                  /note="assembly_name:Contig38"
misc_feature      142500..157956
                  /note="assembly_name:Contig39"
misc_feature      158057..173833
                  /note="assembly_name:Contig40"
BASE COUNT      50115 a 38599 c 38438 g 46468 t 2213 others
ORIGIN

```

```

Query Match      1.28; Score 32; DB 2; Length 175833;
Best Local Similarity 100.0%; Pred. No. 2e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 2703 cccgcacccctccactgagagcgtgtct 2734
      ||||||||||||||||||||||||||||
Db 66179 CCCTGCATCCCTCTCCACTGAGAGCTGTCT 66210

```

```

RESULT 26
AC007380/c      207362 bp      DNA      linear      PRI 29-APR-2000
LOCUS
DEFINITION      Homo sapiens BAC clone RP11-150013 from 2, complete sequence.
ACCESSION      AC007380
VERSION      AC007380.3 GI:5732167
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS      Sulston, J.E. and Waterston, R.
TITLE      Toward a complete human genome sequence
JOURNAL      Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE      99063792
REFERENCE
AUTHORS      Du, F., Maupin, R. and Sowa, D.
TITLE      The sequence of Homo sapiens BAC clone RP11-150013
JOURNAL      Unpublished
REFERENCE
AUTHORS      Waterston, R.H.
TITLE      Direct Submission
JOURNAL      Submitted (25-APR-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 207362)
REFERENCE
AUTHORS      Waterston, R.H.
TITLE      Direct Submission
JOURNAL      Submitted (13-AUG-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 207362)
REFERENCE
AUTHORS      Waterston, R.
TITLE      Direct Submission
JOURNAL      Submitted (29-APR-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Aug 13, 1999 this sequence version replaced gi:5001523.
COMMENT
-----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wustl.wustl.edu
-----
Summary Statistics
-----
Center project name: H_NH0150013

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Feigen, E., Tatem, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)
VECTOR: pBAC3.6
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-355F16. Actual start of this clone is at base position 1 of RP11-150013; actual end is at base position 207362 of RP11-150013.

FEATURES

source

```

1..207362
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-150013"
/clone_1lb="RPCI-11"
13..437
/rpt_family="L1"
repeat_region
438..738
/rpt_family="Alu"
repeat_region
739..1003
/rpt_family="L1"
repeat_region
1160..2723
/rpt_family="L1"
repeat_region
2911..3114
/rpt_family="MERL_type"
repeat_region
3727..3773
/rpt_family="(TTTA)n"
repeat_region
3775..3975
/rpt_family="L2"
repeat_region
5146..5215
/rpt_family="(TA)n"
repeat_region
5848..6147
/rpt_family="Alu"
repeat_region
8227..8472
/rpt_family="Alu"
repeat_region
8876..9591
/rpt_family="L1"
repeat_region
9922..10240
/rpt_family="Other"
repeat_region
11220..11451
/rpt_family="Retroviral"
repeat_region
11784..11949
/rpt_family="MIR"
repeat_region
12040..12079
/rpt_family="L2"
repeat_region
12095..12288

```

```

repeat_region      /rpt_family="Retroviral"
                    12788..13196
repeat_region      /rpt_family="MaLR"
                    13360..13408
repeat_region      /rpt_family="L2"
                    13432..13581
repeat_region      /rpt_family="L2"
                    13624..13921
repeat_region      /rpt_family="Alu"
                    13986..15733
repeat_region      /rpt_family="L1"
                    15734..15777
repeat_region      /rpt_family="Retroviral"
                    15948..17035
repeat_region      /rpt_family="L1"
                    17134..18265
repeat_region      /rpt_family="L1"
                    18932..19021
repeat_region      /rpt_family="MER81"
                    19038..19270
repeat_region      /rpt_family="MIR"
                    19430..20792
repeat_region      /rpt_family="L2"
                    20808..20859
repeat_region      /rpt_family="MER1_type"
                    20889..22563
repeat_region      /rpt_family="L1"
                    25258..25560
repeat_region      /rpt_family="Alu"
                    25642..25793
repeat_region      /rpt_family="MER1_type"
                    25825..25855
repeat_region      /rpt_family="AT_rich"
                    26170..26882
repeat_region      /rpt_family="L1"
                    26883..27043
repeat_region      /rpt_family="Retroviral"
                    27091..27399
repeat_region      /rpt_family="Retroviral"
                    27401..28719
repeat_region      /rpt_family="L1"
                    28720..29042
repeat_region      /rpt_family="Alu"
                    29043..30738
repeat_region      /rpt_family="L1"
                    30777..31443
repeat_region      /rpt_family="L1"
                    31822..32278
repeat_region      /rpt_family="L1"
                    32588..33491
repeat_region      /rpt_family="L2"
                    33500..33681
repeat_region      /rpt_family="MIR"
                    33891..33949
repeat_region      /rpt_family="GA-rich"
                    35130..35210
repeat_region      /rpt_family="MIR"
                    35296..35652
repeat_region      /rpt_family="MaLR"
                    35682..35841
repeat_region      /rpt_family="MIR"
                    35852..36079
repeat_region      /rpt_family="L2"
                    36413..36484
repeat_region      /rpt_family="Mariner"
                    37368..37399
repeat_region      /rpt_family="AT_rich"
                    37607..37780
repeat_region      /rpt_family="Other"
                    37940..38176
repeat_region      /rpt_family="L1"
                    38193..38312
repeat_region      /rpt_family="Other"

```

```

repeat_region      38520..38872
                    /rpt_family="Other"
repeat_region      38873..39166
                    /rpt_family="Alu"
repeat_region      39167..39243
                    /rpt_family="Other"
repeat_region      39244..39290
                    /rpt_family="(TTCA)n"
repeat_region      39291..39494
                    /rpt_family="Other"

Query Match
Best Local Similarity 100.0%; Score 32; DB 9; Length 207362;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2457 actgaattgagacgtctccattggctg 2488
Db 183210 ACTGAATTGAGACCTCTCTCCATTGGCTG 183179

RESULT 27
LOCUS      AL157944      154746 bp      DNA      linear      PRI 19-JAN-2001
DEFINITION Human DNA sequence from clone RP5-1126B2 on chromosome 1p31.1-31.2,
ACCESSION AL157944
VERSION    AL157944.24 GI:12329215
KEYWORDS   HTG.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            Lovell J.
            Direct Submission
            Submitted (19-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
            Requests: clonerequest@sanger.ac.uk
            On Jan 21, 2001 this sequence version replaced gi:12227321.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence has been finished according to sequence map criteria
            as follows. An attempt is made to resolve all sequencing problems,
            such as compressions and repeats, but not necessarily within known
            annotated repeat sequence elements. Where the sequence is
            ambiguous, there is an annotation using the 'unsure' feature key.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            EM: EMBL; SW: SWISSPROT; TR: TrEMBL; WP: WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group. Further information can be found at
            http://www.sanger.ac.uk/HGP/Chrl
            RP5-1126B2 is from the library RPCT-5 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/Bacpac/home.htm
            VECTOR: pCYPAC2
            This sequence is the entire insert of clone RP5-1126B2 The true
            left end of clone RP4-600K12 is at 95393 in this sequence. The true
            left end of clone RP11-170N11 is at 125881 in this sequence. The
            true right end of clone RP11-14M11 is at 52383 in this sequence.
            location/Qualifiers
            1..154746
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="1"
            /map="p31.1-31.2"

```

FEATURES

source

```

misc_feature      /clone="RP5-1126B2"
                  /clone_id="RPCI-5"
                  6468..6551
                  /note="Sequence from overlapping clone bA14M11(AL356974)
                  Assembly confirmed by restriction digest"
misc_feature      65174..65189
                  /note="Single clone region. Assembly confirmed by
                  restriction digest data"
misc_feature      65516..65548
                  /note="Single clone region. Assembly confirmed by
                  restriction digest data"
misc_feature      108104..108297
                  /note="Sequence from overlapping clone dJ600K12(AL357142)
                  Assembly confirmed by restriction digest"
misc_feature      126203..126627
                  /note="Sequence from overlapping clones dJ600K12(AL357142)
                  and bA170N11(AL354654) Assembly confirmed by restriction
                  digest"
misc_feature      126429..126627
                  /note="Sequence from overlapping clones dJ600K12(AL357142)
                  and bA170N11(AL354654) Assembly confirmed by restriction
                  digest"
BASE COUNT      49641 a 27937 c 28135 g 49033 t
ORIGIN

```

```

Query Match      1.1%; Score 31; DB 9; Length 154746;
Best Local Similarity 100.0%; Pred. No. 7.7e-06;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 2462 atgagacctgtctccattgtgtgtct 2492
Db 53581 ATTGAGACCTGCTCCATTGCTGTGCT 53611

```

```

RESULT 28
AC019042/c      168736 bp DNA linear HTG 07-JUL-2000
LOCUS           Homo sapiens chromosome 8 clone RP11-11A18, WORKING DRAFT SEQUENCE,
DEFINITION      5 unordered pieces.
ACCESSION      AC019042
VERSION        AC019042.3 GI:7243881
KEYWORDS       HTG; HTGS_PHASE1; HTGS_DRAFT;
SOURCE         human.
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 168736)
AUTHORS       Waterston, R.H.
TITLE         The sequence of Homo sapiens clone
JOURNAL        Unpublished
AUTHORS       2 (bases 1 to 168736)
TITLE         Waterston, R.H.
JOURNAL        Direct Submission
AUTHORS       Submitted (30-DEC-1999) Genome Sequencing Center, Washington
TITLE         University School of Medicine, 4444 Forest Park Parkway, St. Louis,
JOURNAL        MO 63108, USA
COMMENT        On Mar 15, 2000 this sequence version replaced gi:7139944.

```

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0011A18
----- Summary Statistics -----
Sequencing vector: M13; 77%
Sequencing vector: plasmid; 23%
Chemistry: Dye-terminator Big Dye; 23% of reads
Chemistry: Dye-terminator Big Dye; 23% of reads
Assembly program: Phrap; version 0.99019
Consensus quality: 167158 bases at least Q40
Consensus quality: 167434 bases at least Q30

```

```

Consensus quality: 167583 bases at least Q20
Insert size: 154000; agarose-fp
Insert size: 168336; sum-of-contigs
Quality coverage: 8.04 in Q20 bases; agarose-fp
Quality coverage: 7.38 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 4226: contig of 4226 bp in length
* 4227 4326: gap of unknown length
* 4327 14897: contig of 10571 bp in length
* 14898 14997: gap of unknown length
* 14998 53196: contig of 38199 bp in length
* 53197 53296: gap of unknown length
* 53297 97271: contig of 43975 bp in length
* 97272 97371: gap of unknown length
* 97372 168736: contig of 71365 bp in length.
Location/Qualifiers
1. 168736
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/clone="RP11-11A18"
1. 4226
/note="assembly_name:Contig14"
4327. 14897
/note="assembly_name:Contig15"
14998. 53196
/note="assembly_name:Contig16
clone_end:17
vector_side:left"
53297. 97271
/note="assembly_name:Contig17"
97372. 168736
/note="assembly_name:Contig18"

```

```

BASE COUNT      55308 a 29076 c 28902 g 55038 t 412 others
ORIGIN

```

```

Query Match      1.1%; Score 31; DB 2; Length 168736;
Best Local Similarity 100.0%; Pred. No. 7.6e-06;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 2704 cctgacccctcctccactgagagctgtct 2734
Db 100190 CCTGATCCCTCCTCCACTGAGAGCTGTCT 100160

```

```

RESULT 29
AC024996/c      177864 bp DNA linear PRI 12-DEC-2001
LOCUS           Homo sapiens chromosome 8, clone RP11-697C18, complete sequence.
DEFINITION
ACCESSION      AC024996
VERSION        AC024996.6 GI:17530778
KEYWORDS       HTG.
SOURCE         human.
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 177864)
AUTHORS       Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE         Homo sapiens chromosome 8, clone RP11-697C18
JOURNAL        Unpublished
AUTHORS       2 (bases 1 to 177864)
REFERENCE      Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
AUTHORS       Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
AUTHORS       Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,

```



```

repeat_region complement(14106..14473)
                /rpt_family="THE1B"
repeat_region complement(14513..14922)
                /rpt_family="MSRB"
repeat_region complement(15295..15349)
                /rpt_family="L2"
repeat_region 15350..15369
                /rpt_family="(GAGAA)n"
repeat_region 15991..16303
                /rpt_family="Alusq"
repeat_region complement(16656..16735)
                /rpt_family="MADE1"
repeat_region 17693..17831
                /rpt_family="L2"
repeat_region complement(18046..18221)
                /rpt_family="MER58C"
repeat_region 18594..18614
                /rpt_family="(A)n"
repeat_region 21544..21690
                /rpt_family="MIR"
repeat_region 21741..21946
                /rpt_family="MIR"
repeat_region 22539..22860
                /rpt_family="MER1B"
repeat_region 23154..23175
                /rpt_family="(T)n"
repeat_region 23207..23260
                /rpt_family="MADE1"
repeat_region 23588..23608
                /rpt_family="AT_r1ch"
repeat_region 24080..24101
                /rpt_family="(A)n"
repeat_region complement(24332..24784)
                /rpt_family="MLT1C"
repeat_region 25422..25452
                /rpt_family="AT_r1ch"
repeat_region complement(26347..26643)

```

Query Match 1.1%; Score 31; DB 9; Length 177864;
 Best Local Similarity 100.0%; Pred. No. 7 6e-06;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2704 cctgcacccctccacacagagctgttct 2734
 DB 13356 CCGTCATCCCTCTCCACTGAGAGCTGTCT 13326

RESULT 30
 HSM800283 2328 bp mRNA linear PRI 18-FEB-2000
 LOCUS HSM800283
 DEFINITION Homo sapiens mRNA; cDNA DKFZp564F1422 (from clone DKFZp564F1422).
 ACCESSION AL049942
 VERSION AL049942.1 GI:4884185
 KEYWORDS human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2328)
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Direct Submission
 Submitted (15-MAY-1999) MIPS, Am Klopfersplitz 18a, D-82152
 Martinsried, GERMANY
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;
 sequenced by AGORA (Berlin/Germany) within the CDNA sequencing
 consortium of the German Genome Project.
 This clone (DKFZp564F1422) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further
 information about the clone and the sequencing project is available
 at http://www.mips.biochem.mpg.de/proj/cDNA/.

FEATURES
 Location/Qualifiers

```

source
1..2328
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp564F1422"
/tissue_type="brain"
/clone_lib="564 (synonym: hfb2). Vector pAMP1; host
X1-2blue; sites NotI + SalI"
/dev_stage="fetal"
1..1492
/gene="DKFZp564F1422"
<1..1492
/feature="DKFZp564F1422"
/feature="dub94814.1, strong similarity to Zn finger
proteins"
/codon_start=2
/product="hypothetical protein"
/protein_id="CA843216.1"
/db_xref="GI:4884186"
/translation="RTSGEKPEICVKGCGYTSKSYLVTHERTHTGKPECEGCR
RNDKSSYNKHLKAHSGEKPEVCEGCGFTQSTLVKHOTHSSEKPEVCEGCGRFSO
KSTLVYHQRTHSGEKPEVCEGCGFTQSTLVKHOTHSSEKPEVCEGCGRFSO
TFALHQRTHSGEKPEVCEGCGFTQSTLVKHOTHSSEKPEVCEGCGRFSO
TIHQRTHSGEKPEVCEGCGFTQSTLVKHOTHSSEKPEVCEGCGRFSO
HQRTHSGEKPEVCEGCGFTQSTLVKHOTHSSEKPEVCEGCGRFSO
RTHSGEKPEVCEGCGFTQSTLVKHOTHSSEKPEVCEGCGRFSO
IHRGERPECEGCGRFSO"
2390..2295
polya_signal
2304
BASE COUNT 674 a 501 c 594 g 559 t
ORIGIN

```

Query Match 1.0%; Score 29; DB 9; Length 2328;
 Best Local Similarity 100.0%; Pred. No. 0.00013;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1659 caccagagacacacacagagagagcc 1687
 DB 332 CACCAGAGACACACAGAGAGAGCC 360

RESULT 31
 AK056466 2790 bp mRNA linear PRI 31-OCT-2001
 LOCUS AK056466
 DEFINITION Homo sapiens cDNA FLJ131904 f1s, clone NT2RP7004352.
 ACCESSION AK056466
 VERSION AK056466.1 GI:16551878
 KEYWORDS oligo capping; f1s (full insert sequence).
 SOURCE Homo sapiens testicular carcinoma cell line:NT2 cDNA to mRNA,
 clone lib:NT2RP7 clone:NT2RP7004352.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (sites)
 Tashiro,H., Yamazaki,M., Watanabe,K., Kunagai,A., Itakura,S.,
 Fukushima,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R.,
 Otsuki,T., Sato,H., Wakamatsu,A., Ishi,S., Yamamoto,J., Isono,Y.,
 Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
 Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
 Wagaitsuna,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
 Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
 Nagahari,K., Masuno,Y., Nagai,K. and Isogai,T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 2790)
 Isogai,T., Otsuki,T. and Sugiyama,T.
 Direct Submission
 Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
 Genomics Laboratory; 1532-3 Yana, Kiserazu, Chiba 292-0812, Japan
 (E-mail:genomices@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
 NEDO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES
source
Location/Qualifiers
1..2790
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2R12003220"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP7"
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal precursor cells after 5-weeks retinoic acid (RA) induction."
BASE COUNT
826 a 606 c 712 g 646 t
ORIGIN

Query Match 1.0%; Score 29; DB 9; Length 2790;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1659 caccagagacacacccagggagagacc 1687
|||||
DB 814 CACCAGAGACACACTCAGGGAGAGACC 842

RESULT 32
AK056184 3178 bp mRNA linear PRI 31-OCT-2001
LOCUS Homo sapiens cDNA FLJ131622 fis, clone NT2R12003220, weakly similar
DEFINITION to ZINC FINGER PROTEIN 84.
ACCESSION AK056184
VERSION AK056184.1 GI:16551518
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens teratocarcinoma cell_line:NT2 cDNA to mRNA, clone_lib:NT2R12 clone:NT2R12003220.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS
1 (sites)
Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T., Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahata, K., Masuno, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished

TITLE NEDO human cDNA sequencing project
JOURNAL
AUTHORS
TITLE
JOURNAL
2 (bases 1 to 3178)
Isogai, T., Otsuki, T. and Sugiyama, T.
Direct Submission

COMMENT
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3951, Fax: 81-438-52-3952) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES
source
Location/Qualifiers
1..3178
/organism="Homo sapiens"

/db_xref="taxon:9606"
/clone="NT2R12003220"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2R12"
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal precursor cells treated 2-weeks mitotic inhibitor after 5-weeks retinoic acid (RA) induction. -majorly NT2 neuron"
BASE COUNT
915 a 714 c 831 g 718 t
ORIGIN

Query Match 1.0%; Score 29; DB 9; Length 3178;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1659 caccagagacacacccagggagagacc 1687
|||||
DB 1202 CACCAGAGACACACTCAGGGAGAGACC 1230

RESULT 33
BC021298 3273 bp mRNA linear PRI 22-JAN-2002
LOCUS Homo sapiens, zinc finger protein 337, clone MGC:29686
DEFINITION IMAGE:4551314, mRNA, complete cds.
ACCESSION BC021298
VERSION BC021298.1 GI:18204655
KEYWORDS MGC.

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
AUTHORS
1 (bases 1 to 3273)
Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabps-r@mail.nih.gov
Tissue Procurement: DCTD/DP/Gapzdar
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@ncl.nih.gov
Shenchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Beuffer, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, K., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantipop, S., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov> Series: IRAL Plate: 39 Row: d Column: 14
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

FEATURES
source
Location/Qualifiers
1..3273
/organism="Homo sapiens"
/db_xref="LocusID:26152"
/db_xref="taxon:9606"
/clone="MGC:29686 IMAGE:4551314"
/tissue_type="lung, large cell carcinoma"
/clone_lib="NHL_MGC_18"

CDS

```

/lab_host="DH10B-R"
/notes="Vector: pOT87"
134..2389
/codon_start=1
/product="zinc finger protein 337"
/protein_id="AAH21298.1"
/db_xref="GI:18204656"
/translation="MGPOGARQAFLEAGDVTDFQTEKWRLLSPARALRYETLEN
YSLVSLGILHSKPELIRLEGEVPMGERRRRPGAGITAVEHRLPAPNLIAR
OOLOFSDOSFSDTAEGCEKSTKPMASSPDLRHAVSRHNSVLEESSQORE
NPTEDIKVLGINSRKGAFKACERGDGSRKMMVLIHKASHRQKLFTECHQGR
DESAILLHONTHTGKSYSCVCGRPFSLKANLLRHQTHSGEKPLKVCQGYTSK
SYLVHRTHTGKPYECOCGRFNDKSSYNKHLKASHGEKPFVCEKCGYTSK
FYVAKRIHSEKPYECOCGRFNDKSSYNKHLKASHGEKPFVCEKCGYTSK
RHOHSHGEKPYECOCGRFNDKSSYNKHLKASHGEKPFVCEKCGYTSK
OTHSREKPYECOCGRFNDKSSYNKHLKASHGEKPFVCEKCGYTSK
AHGKERPFECDCGRGFTLAKNLTTHQTHSGEKPFVCEKCGYTSK
SGRPNCKDCGRGFTLAKNLTTHQTHSGEKPFVCEKCGYTSK
KOPVCEKGRGFTLAKNLTTHQTHSGEKPFVCEKCGYTSK
KPEVCEKGRGFTLAKNLTTHQTHSGEKPFVCEKCGYTSK
FCGSGVGEAS"

```

BASE COUNT 988 a 721 c 844 g 720 t

Query Match 1.0%; Score 29; DB 9; Length 3273;
 Best Local Similarity 100.0%; Pred. No. 0.00013;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1659 caccagagacacactcaggaggaagcc 1687
 Db 1229 CACCAGAGACACTCAGGAGGAGACC 1257

RESULT 34
 HS694B14/C
 LOCUS
 DEFINITION Human DNA sequence from clone RP4-694B14 on chromosome 20p11.1-11.22 Contains a novel KRAB box protein with 18 C2H2 type zinc finger domains, a novel haloacid dehalogenase-like hydrolase family protein similar to (archaea) bacterial proteases, two putative novel genes, a novel pseudogene, ESTs, an STS, GSSs and three CpG islands, complete sequence.

ACCESSION AL031673 GI:11968366
 VERSION AL031673.19
 KEYWORDS HTG; Cpg island; haloacid dehalogenase; hydrolase; KRAB box; zinc finger.

SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 127867)
 AUTHORS Smith, M.
 TITLE Direct Submission
 JOURNAL Submitted (20-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

COMMENT

On Dec 22, 2000 this sequence version replaced gi:10198607.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
 Bm: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WORMPEP database can be found at
<http://www.sanger.ac.uk/projects/C.elegans/wormep>
 This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr20>

RP4-694B14 is from the library RPI-4 constructed by the group of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pCYPAC2
 This sequence is the entire insert of clone RP4-694B14. The true left end of clone RP13-401N8 is at 116161 in this sequence. The true right end of clone RP11-9616 is at 20586 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

FEATURES

FEATURE	location/Qualifiers
repeat_region	1..127867
repeat_region	/organism="Homo sapiens"
repeat_region	/db_xref="taxon:9606"
repeat_region	/chromosome="20"
repeat_region	/map="p11.1-11.22"
repeat_region	/clone="RP4-694B14"
repeat_region	/clone="lib="RPI-4"
repeat_region	1..74
repeat_region	/note="Aluub repeat: matches 194..266 of consensus"
repeat_region	787..1088
repeat_region	/note="AluSp repeat: matches 1..303 of consensus"
repeat_region	1513..1786
repeat_region	/note="AluX repeat: matches 1..295 of consensus"
repeat_region	1838..2372
repeat_region	/note="L1MC3 repeat: matches 7230..7739 of consensus"
repeat_region	2373..2669
repeat_region	/note="Aluub repeat: matches 1..298 of consensus"
repeat_region	2670..3427
repeat_region	/note="L1MC3 repeat: matches 6508..7230 of consensus"
repeat_region	3428..3725
repeat_region	/note="Aluub repeat: matches 1..298 of consensus"
repeat_region	3726..4094
repeat_region	/note="L1MC3 repeat: matches 6157..6508 of consensus"
repeat_region	4095..4449
repeat_region	/note="THE1A repeat: matches 1..354 of consensus"
repeat_region	4450..4592
repeat_region	/note="L1MC3 repeat: matches 6022..6157 of consensus"
repeat_region	4593..4898
repeat_region	/note="AluSc repeat: matches 1..306 of consensus"
repeat_region	4899..5002
repeat_region	/note="L1MC3 repeat: matches 5924..6022 of consensus"
repeat_region	5003..5309
repeat_region	/note="Aluuo repeat: matches 1..302 of consensus"
repeat_region	5310..5576
repeat_region	/note="L1MC3 repeat: matches 5460..5924 of consensus"
repeat_region	5904..6210
repeat_region	/note="AluX repeat: matches 1..307 of consensus"
repeat_region	6481..6579
repeat_region	/note="L2 repeat: matches 2634..2743 of consensus"
repeat_region	6626..7016
repeat_region	/note="MLT1B repeat: matches 6..390 of consensus"
repeat_region	7565..7652
repeat_region	/note="44 copies 2 mer ct 73% conserved"
repeat_region	7668..7781
repeat_region	/note="57 copies 2 mer ct 75% conserved"
repeat_region	complement(7844..8286)
repeat_region	/note="Random repeat: region contains forced join and single clone. HindIII digest suggests there is approximately 1kb missing from the tandem repeat."
repeat_region	8095..8132
repeat_region	/note="19 copies 2 mer ca 78% conserved"
repeat_region	complement(8206..8578)
repeat_region	/note="match: GSS: Em:A0535283"
repeat_region	complement(8211..8572)
repeat_region	/note="match: GSS: Em:A0507005"
repeat_region	complement(8265..8579)
repeat_region	/note="match: GSS: Em:A0587567"
repeat_region	8822..9429

repeat__region	/note="OpG island" /evidence=not_experimental 10072. .10203	/note="FLAM_C repeat: matches 2. .133 of consensus"
repeat__region	/note="FLAM_C repeat: matches 2. .133 of consensus" 10737. .11028	/note="Aluux repeat: matches 6. .298 of consensus"
repeat__region	11364. .11671	/note="Aluux repeat: matches 1. .304 of consensus"
repeat__region	12208. .12496	/note="Aluub repeat: matches 1. .299 of consensus"
repeat__region	12508. .12813	/note="Aluub repeat: matches 1. .301 of consensus"
repeat__region	13209. .13290	/note="Aluub repeat: matches 1. .301 of consensus"
repeat__region	13319. .13366	/note="MLTIF repeat: matches 243. .317 of consensus"
repeat__region	13394. .13480	/note="MLTIF repeat: matches 280. .327 of consensus"
repeat__region	13501. .13636	/note="MLTIF repeat: matches 468. .547 of consensus"
repeat__region	13828. .14126	/note="FLAM_C repeat: matches 1. .133 of consensus"
misc__feature	/note="Aluub repeat: matches 1. .304 of consensus" complement(13329. .14330)	
repeat__region	14518. .14584	/note="match: GSS: Em:AQ375915"
repeat__region	/note="L2 repeat: matches 2677. .2745 of consensus" 14531. .14757	
repeat__region	/note="L2 repeat: matches 1709. .1875 of consensus" 14937. .15009	
misc__feature	/note="Aluub/FLAM_C repeat: matches 1. .73 of consensus" complement(15027. .15331)	
misc__feature	/note="match: GSS: Em:AQ280927" complement(15040. .15332)	
repeat__region	/note="match: STRs: Em:HS528L19S" 15040. .15332	/note="Aluux repeat: matches 1. .293 of consensus"
misc__feature	complement(15061. .15328)	
misc__feature	/note="match: GSS: Em:A518424" complement(15065. .15333)	
misc__feature	/note="match: GSS: Em:AQ74659S" complement(15066. .15246)	
misc__feature	/note="match: GSS: Em:AQ807616" complement(15089. .15356)	
misc__feature	/note="match: GSS: Em:A490556" 15091. .15333	
misc__feature	/note="match: STRs: Em:G20204" complement(15097. .15332)	
misc__feature	/note="match: GSS: Em:A521752" 15109. .15332	
misc__feature	/note="match: GSS: Em:B01680" 15111. .15332	
misc__feature	/note="match: GSS: Em:A521733" 15114. .15332	
misc__feature	/note="match: GSS: Em:AQ352998" 15117. .15332	
misc__feature	/note="match: GSS: Em:A519283" 15121. .15331	
misc__feature	/note="match: GSS: Em:A520451" 15121. .15332	
misc__feature	/note="match: GSS: Em:AQ223482 Em:AQ631465" complement(15127. .15356)	
misc__feature	/note="match: GSS: Em:AQ115261" 15129. .15340	
misc__feature	/note="match: GSS: Em:AQ532262" complement(15135. .15332)	
misc__feature	/note="match: GSS: Em:AQ618206" 15143. .15329	
misc__feature	/note="match: GSS: Em:AQ791131" 15156. .15348	
misc__feature	/note="match: GSS: Em:AQ536385" complement(15160. .15332)	
misc__feature	/note="match: STRs: Em:G42878" complement(15345)	

```

misc_feature /note="match: GSS: Em:AQ245808  
complement(15171..15344)"  
/note="match: STS: Em:G15778"  
misc_feature /note="match: GSS: Em:AQ315996"  
complement(15182..15328)  
/note="match: GSS: Em:AQ315996"  
complement(15188..15332)  
/note="match: STS: Em:G07766"  
15192..15332  
misc_feature /note="match: GSS: Em:AQ279850"  
complement(15194..15356)  
/note="match: GSS: Em:AQ081971"  
complement(15200..15333)  
/note="match: GSS: Em:B88705"  
15207..15331  
/note="match: STS: Em:HS198VD10"  
15216..15353  
/note="match: GSS: Em:AG014867"  
Query Match 1.0%; Score 29; DB 9; Length 127867;  
Best Local Similarity 100.0%; Pred. No. 0.00011;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1659 caccagagacacactcagggagaagcc 1687  
|||||caccagagacacactcagggagaagcc 1687  
DB 82604 CACCAGAGACACACTCAGGGAGAAGCC 82576  
RESULT 35  
AC013770  
LOCUS AC013770 170129 bp DNA linear HTG 22-MAR-2000  
DEFINITION Homo sapiens clone RP11-3B10, WORKING DRAFT SEQUENCE, 15 unordered  
pieces.  
AC013770  
AC013770.3 GI:7280299  
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 170129)  
Bairren,B., Linton,L., Nusbaum,C. and Lander,E.  
Homo sapiens, clone RP11-3B10  
Unpublished  
2 (bases 1 to 170129)  
Bairren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
Baldwin,J., Barna,N., Beckery,R., Boguslavsky,L., Boukhalter,B.,  
Brown,A., Castle,A., Colangelo,M., Collins,S., Collamore,A.,  
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donejan,L., Doyle,M.,  
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,  
Galland,J.C., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
Lehocky,J., Liu,C., Locke,K., Macdonald,P., Marguis,N.,  
McEwan,P., McGurt,A., McKernan,K., McLaughlin,J., Meldrim,J.,  
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Testafte,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (15-NOV-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 22, 2000 this sequence version replaced gi:7209490.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence.submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L2702  
Center clone name: 3_B-10

```

misc_feature	15558.	.23562	/note="assembly_fragment"	
misc_feature	23663.	31772	/note="assembly_fragment"	
misc_feature	31873.	.43485	/note="assembly_fragment"	
misc_feature	43586.	.55357	/note="assembly_fragment"	
misc_feature	55458.	.66468	/note="assembly_fragment"	
misc_feature	66569.	.95233	/note="assembly_fragment"	
misc_feature	95334.	.124230	/note="assembly_fragment"	
misc_feature	12431.	.170129	/note="assembly_fragment"	
BASE COUNT	53001	a 31892	c 33160	g 50668 t 1408 others
ORIGIN				

Query Match	1.0%;	Score 29;	DB 2;	Length 170129;
Best Local Similarity	100.0%;	Prod. No. 0.00011;		
Matches	29;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	2470	ccgtgcttcccatctgtgtgtccttcctc	2498	
Db	33385	CCTGCTCTCCCATTTGGTGTCTTTCC	33413	
RESULT	36			
AC109473/c		181770 bp	DNA	linear
LOCUS				HTG 04-FEB-2002
DEFINITION		Homo sapiens chromosome 5 clone RPL1-3B10,		WORKING DRAFT SEQUENCE,
		9 unnumbered pieces.		
AC109473				
AC109473.1		GI:18483469		
VERSION				
KEYWORDS		HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFN.		
SOURCE		human.		
ORGANISM		Homo sapiens		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE		1 (bases 1 to 181770)		
AUTHORS		DOE Joint Genome Institute.		
TITLE		Sequencing of Human Chromosome 5		
JOURNAL		Unpublished		
REFERENCE		2 (bases 1 to 181770)		
AUTHORS		DOE Joint Genome Institute.		
TITLE		Direct Submission		
JOURNAL		Submitted (04-FEB-2002)		
		Production Sequencing Facility, DOE Joint		

```

COMMENT
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web Site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 407551
Center Clone name: RPC1-11_3B10
-----
Summary Statistics
Consensus quality: 174095 bases at least Q40
Consensus quality: 175269 bases at least Q30
Consensus quality: 176156 bases at least Q20
Estimated insert size: 174280: agarose-fp estimation
Estimated insert size: 180970: sum-of-contigs estimation
Quality coverage: 43.3 In Q20 bases: agarose-fp estimation
Quality coverage: 41.7 In Q20 bases: sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

```

* as soon as it is available and the accession number will
 * be preserved.
 * 1
 * 1325: contig of 1325 bp in length
 * 1425: gap of unknown length
 * 1426: contig of 1146 bp in length
 * 2572: gap of unknown length
 * 2672: contig of 1361 bp in length
 * 4033: gap of unknown length
 * 4133: contig of 1060 bp in length
 * 5193: gap of unknown length
 * 5293: contig of 1241 bp in length
 * 6534: gap of unknown length
 * 6634: contig of 2034 bp in length
 * 8668: gap of unknown length
 * 8767: contig of 1198 bp in length
 * 9666: gap of unknown length
 * 10065: contig of 1237 bp in length
 * 11303: gap of unknown length
 * 11403: contig of 170368 bp in length.
 Location/Qualifiers
 1. 181770
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="RP11-3810"
 /clone_lib="RPC1 human BAC library 11"

BASE COUNT 57236 a 34865 c 35014 g 53844 t 811 others
 ORIGIN

Query Match 1.0%; Score 29; DB 2; Length 181770;
 Best Local Similarity 100.0%; Pred. No. 0.00011;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2470 cctgtcttccattggtgtgcttctc 2498
 ||||||||||||||||||||||||||||
 Db 177072 CCTGCTCCCAATTGTGTCTTCCTC 177044

RESULT 37
 BC011862 2823 bp mRNA linear PRI 02-ANG-2001
 LOCUS Homo sapiens, clone MGC:20504 IMAGE:4328382, mRNA, complete cds.
 ACCESSION BC011862
 VERSION BC011862.1 GI:15080179
 KEYWORDS MGC.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 2823)
 Direct Submission
 Submitted (30-JUL-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nigri.nih.gov
 Shevchenko, Y., Weetherby, K.D., Beckstrom-Sternberg, S.M.,
 Benjamini, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
 Dierick, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
 Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Snyder, B., Stantrilpop, S., Thomas, P.D.,

REMARK
 COMMENT

Tiongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
 Zhang, L.-H. and Green, E.D.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL ac: <http://image.llnl.gov>
 Series: IRML Plate: 28 Row: m Column: 19
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Genomescan gene
 prediction.

FEATURES
 source Location/Qualifiers
 1. 2823
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:20504 IMAGE:4328382"
 /tissue_type="uterus, leiomyosarcoma"
 /clone_lib="NIH-MGC_46"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 292..1467
 /codon_start=1
 /product="Unknown (protein for MGC:20504)"
 /protein_id="AAH11862.1"
 /db_xref="GI:15080180"
 /translation="MVFETPSSAQRPNPVQIDKGLKETLRFCAINCREYEPDHL
 ESNFTNPTLLGKPKYICSDGRSFKDSTLIRHRIHSMKRPVYCEGSGFSK
 NLSRQRTHESEKPYLCREGCSFRSKSLNHWTHSEKPYVCEGSGFSKSF
 IRRHRTSGSEKPYVLCREGCSFRSKSLNHWTHSEKPYVCEGSGFSKSLDIK
 HORTHLDEKPYVCEGSGFSKSLNHWTHSEKPYVCEGSGFSKSLIYVHRT
 RHSGEKPYVCEGSGFSKSLNHWTHSEKPYVCEGSGFSKSLIYVHRT
 HSGEKPYVCEGSGFSKSLNHWTHSEKPYVCEGSGFSKSLIYVHRT"

BASE COUNT 812 a 638 c 749 g 624 t
 ORIGIN

Query Match 1.0%; Score 28; DB 9; Length 2823;
 Best Local Similarity 100.0%; Pred. No. 0.00049;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1659 caccagagacacacacacaggaagc 1686
 ||||||||||||||||||||||||
 Db 1114 CACCAGAGACACACTAGGGAGAGAC 1141

RESULT 38
 AK05734 3055 bp mRNA linear PRI 31-OCT-2001
 LOCUS Homo sapiens CDNA FLJ31172 fis, clone KIDNE2000051, moderately
 DEFINITION similar to ZINC FINGER PROTEIN 133.
 ACCESSION AK05734
 VERSION AK05734.1 GI:16550536
 KEYWORDS oligo capping; fls (full insert sequence).
 SOURCE Homo sapiens kidney CDNA to mRNA, clone_11b:KIDNE2
 clone:KIDNE2000051.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (sites)
 Ishibashi, T., Kanehori, K., Yosida, M., Watanabe, S., Ishida, S.,
 Ono, Y., Horita, T., Hiraoaka, S., Murakawa, K., Takiguchi, S.,
 Kusano, Y., Watanabe, M., Fujimori, K., Tanai, H., Ishida, M.,
 Yamashita, H., Chiba, Y., Sugiyama, T., Irie, R., Otsuki, T., Sato, H.,
 Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y.,
 Saito, K., Nishikawa, T., Kimura, K., Matsuo, K., Nakamura, Y.,
 Sekine, M., Kikuchi, H., Kanda, K., Magatsuna, M., Takahashi-Fujii, A.,
 Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
 Nagahari, K., Masuno, Y., Nagai, K. and Isogai, T.
 NEDO human CDNA sequencing project
 Unpublished
 2 (bases 1 to 3055)
 Isogai, T., Otsuki, T. and Sugiyama, T.
 Direct Submission
 Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
 Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT (E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan: cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'-6'3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: RAB and
HRI.

FEATURES
source location/Qualifiers
1..3055
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="KIDNE2000051"
/tissue_type="kidney"
/clone_1lb="KIDNE2"
/note="cloning vector: pME18SFL3"
188..1171
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71000.1"
/db_xref="gi:16550537"
/translation="MLENRYNLSIAEPKPEITYTSCSLAFSCQFLSHVQVIFIG
LCAENHFPENSSPGHKHKOQGYSHVSCFENAEGERGSGSPMSARTEERTISA
FSPYLOKQASPRKKNVETEPSAQRPNPVDLKGKLETLRGALICREYEDPD
NLSNFIITNPTLLGKRPICSDCGRSFKDSTLIRHHRISHMEKPYVCGECRGESQ
KSNLSNRTORTSEKPYLCRECGSFKSKILNRHWTSEKPYVCGECRGESQ
SFIRHRTSGEKPVCYCEGRCGFCDSKTLKHORHSGEKPYVCGECRGESQNSLI
IKHQRTHDEKPYVCGECRGFCPCDSTLIIHRTSGEKPYVCGECRGESQSLIIV
HQRTHSGEKPYVCGECRGFSQSKNLIRHRTSNEKPYLCRECGRCFCRCKSLIIVH
RHSGERPYVCGECRGFSKSLILVHQRTHSGEKPYVCGECRGFSKSLIRHQRTH

BASE COUNT 866 a 692 c 808 g 689 t

ORIGIN

Query Match 1.0%; Score 28; DB 9; Length 3055;
Best Local Similarity 100.0%; Pred. No. 0.00049;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1659 caccagagagacacactcaggaggaagc 1686
|||||
Db 1364 CACCAGAGGACACACTCAGGGAGAGC 1391

RESULT 39 61452 bp DNA linear HTG 09-JUN-2001
AC091936
LOCUS Homo sapiens chromosome 5 clone RP11-325C20, WORKING DRAFT
DEFINITION
AC091936
SEQUENCE 16 unordered pieces.
AC091936
VERSION AC091936.1 GI:14333872
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 61452)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 5
Unpublished
2 (bases 1 to 61452)
DOE Joint Genome Institute.
Direct Submission
Submitted (09-JUN-2001) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Project Name: 531233

Center clone name: RPCT-11_325C20

Summary Statistics
Consensus quality: 47683 bases at least Q40
Consensus quality: 53520 bases at least Q30
Consensus quality: 56595 bases at least Q20
Estimated insert size: 160000; agarose-fp estimation
Estimated insert size: 59952; sum-of-contigs estimation
Quality coverage: 5.17 in Q20 bases; agarose-fp estimation
Quality coverage: 13.8 in Q20 bases; sum-of-contigs estimation.
NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1158: contig of 1158 bp in length
* 1159 1258: gap of unknown length
* 1259 2621: contig of 1363 bp in length
* 2622 2721: gap of unknown length
* 2722 3952: contig of 1231 bp in length
* 3953 4052: gap of unknown length
* 4053 5251: contig of 1199 bp in length
* 5252 5351: gap of unknown length
* 5352 6514: contig of 1163 bp in length
* 6515 6614: gap of unknown length
* 6615 7791: contig of 1177 bp in length
* 7792 7891: gap of unknown length
* 7892 9549: contig of 1658 bp in length
* 9550 9650: gap of unknown length
* 9650 11835: contig of 2186 bp in length
* 11836 13920: gap of unknown length
* 13921 14020: contig of 1985 bp in length
* 14021 15094: gap of unknown length
* 15095 15194: gap of 1074 bp in length
* 15195 16490: contig of 1296 bp in length
* 16491 16590: gap of unknown length
* 16591 19279: contig of 2689 bp in length
* 19280 19379: gap of unknown length
* 19380 29586: contig of 10207 bp in length
* 29587 29686: gap of unknown length
* 29687 39306: contig of 9620 bp in length
* 39307 39406: gap of unknown length
* 39407 39407: gap of unknown length
* 39407 48479: contig of 9072 bp in length
* 48479 48578: gap of unknown length
* 48579 61452: contig of 12874 bp in length.

FEATURES
source location/Qualifiers
1..61452
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-325C20"
/clone_1lb="RPCT-11_325C20"
/clone_1lb="RPCT-11_325C20"

BASE COUNT 18087 a 10746 c 10505 g 20606 t 1508 others

ORIGIN

Query Match 1.0%; Score 28; DB 2; Length 61452;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2501 atgattcccaaccctcattattac 2528
|||||
Db 22636 ATTGATCCCAACCTTACCTATTATAC 22663

RESULT 40 62002 bp DNA linear HTG 13-JUN-2001
AF161800
LOCUS Homo sapiens chromosome 8 clone XX-189m5 map 8q21.2, *** SEQUENCING
DEFINITION


```

* 1263 12967: contig of 705 bp in length
* 12968 13067: gap of 100 bp
* 13068 13808: contig of 741 bp in length
* 13809 13908: gap of 100 bp
* 13909 14621: contig of 713 bp in length
* 14622 14721: gap of 100 bp
* 14722 15434: contig of 713 bp in length
* 15435 15534: gap of 100 bp
* 15535 16257: contig of 723 bp in length
* 16258 16357: gap of 100 bp
* 16358 17074: contig of 717 bp in length
* 17075 17174: gap of 100 bp
* 17175 17901: contig of 727 bp in length
* 17902 18001: gap of 100 bp
* 18002 18716: contig of 715 bp in length
* 18717 18816: gap of 100 bp
* 18817 19523: contig of 707 bp in length
* 19524 19623: gap of 100 bp
* 19624 20307: contig of 684 bp in length
* 20308 20407: gap of 100 bp
* 20408 21109: contig of 702 bp in length
* 21110 21209: gap of 100 bp
* 21210 21913: contig of 704 bp in length
* 21914 22013: gap of 100 bp
* 22014 22735: contig of 722 bp in length
* 22736 22835: gap of 100 bp
* 22836 23556: contig of 721 bp in length
* 23557 23656: gap of 100 bp
* 23657 24379: contig of 723 bp in length
* 24380 24479: gap of 100 bp
* 24480 25134: contig of 655 bp in length
* 25135 25234: gap of 100 bp
* 25235 25962: contig of 728 bp in length
* 25963 26062: gap of 100 bp
* 26063 26803: contig of 741 bp in length
* 26804 26903: gap of 100 bp
* 26904 27638: contig of 735 bp in length
* 27639 27738: gap of 100 bp
* 27739 28427: contig of 689 bp in length
* 28428 28527: gap of 100 bp
* 28528 29224: contig of 697 bp in length
* 29225 29324: gap of 100 bp
* 29325 30033: contig of 709 bp in length
* 30034 30133: gap of 100 bp
* 30134 30836: contig of 703 bp in length
* 30837 30936: gap of 100 bp
* 30937 31636: contig of 700 bp in length
* 31637 31736: gap of 100 bp
* 31737 32481: contig of 745 bp in length
* 32482 32581: gap of 100 bp
* 32582 33287: contig of 706 bp in length
* 33288 33387: gap of 100 bp
* 33388 34096: contig of 709 bp in length
* 34097 34196: gap of 100 bp
* 34197 34915: contig of 719 bp in length
* 34916 35015: gap of 100 bp
* 35016 35737: contig of 722 bp in length
* 35738 35837: gap of 100 bp
* 35838 36542: contig of 705 bp in length
* 36543 36642: gap of 100 bp
* 36643 37357: contig of 715 bp in length
* 37358 37457: gap of 100 bp
* 37458 38146: contig of 689 bp in length
* 38147 38246: gap of 100 bp
* 38247 38966: contig of 720 bp in length
* 38967 39066: gap of 100 bp
* 39067 39788: contig of 722 bp in length
* 39789 39888: gap of 100 bp
* 39889 40623: contig of 735 bp in length
* 40624 40723: gap of 100 bp
* 40724 41446: contig of 723 bp in length
* 41447 41546: gap of 100 bp
* 41547 42259: contig of 713 bp in length

```

```

* 42260 42359: gap of 100 bp
* 42360 43092: contig of 733 bp in length
* 43093 43192: gap of 100 bp
* 43193 43918: contig of 726 bp in length
* 43919 44018: gap of 100 bp
* 44019 44746: contig of 728 bp in length
* 44747 44846: gap of 100 bp
* 44847 45554: contig of 708 bp in length
* 45555 45654: gap of 100 bp
* 45655 46368: contig of 714 bp in length
* 46369 46468: gap of 100 bp
* 46469 47197: contig of 729 bp in length
* 47198 47297: gap of 100 bp
* 47298 48020: contig of 723 bp in length
* 48021 48120: gap of 100 bp
* 48121 48851: contig of 731 bp in length
* 48852 48951: gap of 100 bp
* 48952 49667: contig of 716 bp in length
* 49668 49767: gap of 100 bp
* 49768 50502: contig of 735 bp in length
* 50503 50602: gap of 100 bp
* 50603 51289: contig of 687 bp in length
* 51290 51389: gap of 100 bp
* 51390 52124: contig of 735 bp in length
* 52125 52224: gap of 100 bp
* 52225 52885: contig of 661 bp in length
* 52886 52985: gap of 100 bp
* 52986 53698: contig of 713 bp in length
* 53699 53798: gap of 100 bp
* 53799 54520: contig of 722 bp in length
* 54521 54620: gap of 100 bp
* 54621 55337: contig of 717 bp in length
* 55338 55437: gap of 100 bp
* 55438 56170: contig of 733 bp in length

```

Query Match 1.0%: Score 28; DB 2; Length 65236;
 Best Local Similarity 100.0%: Pred. No. 0.00045;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2461 aattgaagacctgtctccattggtg 2488
 Db 3609 AATTGAGACCTGTCTTCCATTGGTCG 3636

RESULT 42
 AL512489
 LOCUS
 DEFINITION Human DNA sequence from clone RP11-44708 on chromosome 6, complete
 sequence.
 ACCESSION AL512489
 VERSION AL512489.11 GI:13992076
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 80846)
 REFERENCE
 AUTHORS Hammond, S.
 TITLE Direct Substitution
 JOURNAL Submitted (04-MAY-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 requests: clonerequests@sanger.ac.uk
 On May 8, 2001 this sequence version replaced gi:13751525.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL, Sw, SWSISPPROT, Tr, TREMBL, Wp, WORMEP. Information on the WORMEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/Wormep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6> Rpl1-44708 is from the library RPl-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone Rpl1-44708 it may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RPl-162C6 is at 80747 in this sequence. The true right end of clone RPl-271G9 is at 100 in this sequence.

FEATURES

SOURCE

```
1. 80846
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="Rpl1-44708"
/clone.lib="RPl-11.2"
1. 268
/note="L1P repeat: matches 3697. .3964 of consensus"
repeat_region
380. .448
/note="3 copies 23 mer 82% conserved"
repeat_region
430. .473
/note="11 copies 4 mer tagg 86% conserved"
repeat_region
506. .595
/note="45 copies 2 mer ta 62% conserved"
repeat_region
508. .591
/note="21 copies 4 mer taga 84% conserved"
repeat_region
2590. .2623
/note="17 copies 2 mer tg 85% conserved"
repeat_region
3664. .3969
/note="L1MB8 repeat: matches 5847. .6173 of consensus"
repeat_region
4687. .4765
/note="L2 repeat: matches 2098. .2176 of consensus"
repeat_region
5352. .5419
/note="34 copies 2 mer at 75% conserved"
repeat_region
6233. .6631
/note="Charliela repeat: matches 1025. .1455 of consensus"
repeat_region
6632. .7214
/note="L1R28 repeat: matches 1. .608 of consensus"
repeat_region
7279. .7323
/note="MER52A repeat: matches 1207. .1252 of consensus"
repeat_region
7361. .7515
/note="MER52A repeat: matches 1602. .1750 of consensus"
repeat_region
7543. .8513
/note="Charliela repeat: matches 6. .1022 of consensus"
repeat_region
8659. .8961
/note="Alusg repeat: matches 1. .302 of consensus"
misc_feature
complement(9091. .9549)
/note="match: GSS: Em:AQ728035"
repeat_region
9294. .9423
/note="FLAM_C repeat: matches 1. .131 of consensus"
misc_feature
9582. .10277
/note="match: GSS: Em:AQ626592"
repeat_region
11212. .11251
/note="20 copies 2 mer tt 87% conserved"
repeat_region
12790. .12897
/note="MIR repeat: matches 118. .232 of consensus"
misc_feature
complement(12908. .13388)
/note="match: GSS: Em:AQ437578"
repeat_region
13359. .13420
/note="MIR repeat: matches 17. .78 of consensus"
misc_feature
13377. .14279
```

```
repeat_region
14124. .14355
/note="match: GSS: Em:AQ748923"
repeat_region
14357. .14562
/note="L2 repeat: matches 2514. .2748 of consensus"
repeat_region
15005. .15315
/note="L1MA2 repeat: matches 6087. .6308 of consensus"
repeat_region
15710. .15953
/note="AluY repeat: matches 1. .308 of consensus"
repeat_region
16730. .17022
/note="MIR repeat: matches 5. .262 of consensus"
misc_feature
complement(17405. .17893)
/note="AlusX repeat: matches 1. .291 of consensus"
repeat_region
18033. .18161
/note="match: GSS: Em:AQ672286"
repeat_region
20356. .20542
/note="MIR repeat: matches 23. .153 of consensus"
repeat_region
20701. .20897
/note="MIR repeat: matches 2. .189 of consensus"
repeat_region
21189. .21273
/note="MER63A repeat: matches 1. .200 of consensus"
repeat_region
21514. .21663
/note="L2 repeat: matches 2624. .2710 of consensus"
repeat_region
21797. .22062
/note="MIR repeat: matches 77. .231 of consensus"
repeat_region
22063. .22700
/note="L1PA6 repeat: matches 5233. .5497 of consensus"
misc_feature
23497. .23826
/note="L1PA6 repeat: matches 5497. .6141 of consensus"
misc_feature
23505. .23963
/note="match: GSS: Em:B80555"
repeat_region
23934. .23959
/note="match: GSS: Em:AQ235419"
repeat_region
25593. .31056
/note="L3 copies 2 mer aa 100% conserved"
repeat_region
31044. .32076
/note="L1PB1 repeat: matches -279. .6155 of consensus"
repeat_region
32085. .32551
/note="L1PB1 repeat: matches -1537. -.412 of consensus"
repeat_region
32852. .32907
/note="Tigger2a repeat: matches 2. .434 of consensus"
repeat_region
32872. .32907
/note="28 copies 2 mer tg 76% conserved"
misc_feature
complement(33777. .34116)
/note="9 copies 4 mer tgtg 91% conserved"
misc_feature
34248. .34564
/note="match: GSS: Em:A2246003"
misc_feature
complement(35522. .36045)
/note="match: GSS: Em:A246003"
repeat_region
35729. .41433
/note="match: GSS: Em:A0791133"
repeat_region
41074. .41617
/note="L1PA7 repeat: matches 549. .6140 of consensus"
repeat_region
42001. .42188
/note="match: GSS: Em:AQ231043"
repeat_region
42244. .42338
/note="L2 repeat: matches 2548. .2741 of consensus"
repeat_region
42374. .42653
/note="MER5A repeat: matches 13. .105 of consensus"
repeat_region
43032. .43265
/note="L2 repeat: matches 2246. .2519 of consensus"
repeat_region
43855. .44065
/note="MIR repeat: matches 13. .262 of consensus"
repeat_region
44137. .44860
/note="M1T1J repeat: matches 293. .516 of consensus"
repeat_region
44930. .44967
/note="L1PA7 repeat: matches 5405. .6141 of consensus"
repeat_region
45003. .45194
/note="19 copies 2 mer ta 78% conserved"
repeat_region
45135. .45194
/note="96 copies 2 mer at 58% conserved"
repeat_region
45200. .45237
/note="15 copies 4 mer atat 75% conserved"
repeat_region
/note="19 copies 2 mer at 78% conserved"
```

```
repeat_region 46224..46423
                /note="AluSg/x repeat: matches 113. .312 of consensus"
repeat_region 46477..46617
                /note="HERVL repeat: matches 3164. .3304 of consensus"
repeat_region 47132..47713
                /note="L1MEC repeat: matches 1392. .1977 of consensus"
repeat_region 47950..47991
                /note="L1MEC repeat: matches 2272. .2313 of consensus"
repeat_region 48550..48762
                /note="L1ME1 repeat: matches 5956. .6168 of consensus"
misc_feature 50034..50527
                /note="match: GSS: Em:AQ732908"
repeat_region 50281..51058
                /note="L1PA2 repeat: matches 4. .776 of consensus"
repeat_region 51054..516274
                /note="L1PA2 repeat: matches 900. .6146 of consensus"
repeat_region 56466..56574
                /note="AluSg/x repeat: matches 1. .109 of consensus"
repeat_region 56577..56785
                /note="MLT2 repeat: matches 117. .337 of consensus"
repeat_region 57436..57706
                /note="AluSg repeat: matches 5. .287 of consensus"
repeat_region 59298..60448
                /note="L1ME repeat: matches 890. .1593 of consensus"
repeat_region 60497..60587
                /note="L1ME repeat: matches 2128. .2214 of consensus"
repeat_region 60588..60904
```

```
Query Match 1.0%; Score 28; DB 9; Length 80846;
Best Local Similarity 100.0%; Pred. No. 0.00044;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2358 tgatagtgcagagagcagcgaatgcc 2385
Db 6632 TGATAGTGCAGAGGAGCAATGCC 6659
```

```
RESULT 43
HSJ734P14
LOCUS
DEFINITION
```

HSJ734P14 90220 bp. DNA linear PRI 23-FEB-2001
Human DNA sequence from clone RP4-734P14 on chromosome 20. Contains
the 3' end of a novel KRAB box protein gene, a novel C2H2 zinc
finger domain protein gene, the SNRPB gene for small nuclear
ribonucleoprotein polypeptides B and B1 (B') and the 3' end of the
gene for a novel transglutaminase. Contains ESTs, STSs, GSSs and a
CpG island, complete sequence.

```
ACCESSION AL049650 GI:5123801
VERSION AL049650.8 GI:5123801
KEYWORDS HTG; CpG island; KRAB box; snRNP; SNRPB; transglutaminase; Zinc  
finger.
```

```
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 90220)  
Blakey, S.  
Direct Submission  
Submitted (22-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
requests: clonerequest@sanger.ac.uk
```

COMMENT

On Jun 22, 1999 this sequence version replaced gi:4995635.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; SW:
SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP
database can be found at
<http://www.sanger.ac.uk/projects/Celegans/wormpep> This sequence
was generated from part of bacterial clone contigs of human
chromosome 20, constructed by the Sanger Centre Chromosome 20
Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr20>
RP4-734P14 is from the library RP4-4 constructed by the group of
Pleier de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
VECTOR: pCYPAC2
IMPORTANT: This sequence is not the entire insert of clone
RP4-734P14. It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true left end of clone RP4-734P14 is at 1 in this sequence.
The true left end of clone RP4-816K17 is at 90117 in this sequence.

FEATURES

source

repeat_region

mRNA

gene

CDS

```
gene
    join(929..1046,1247..1305,1999..2125,2740..2792)
    /gene="dJ734P14.1"
    /note="protein homologs suggest that this gene and gene  
5 are part of the same gene in the form of a yet to be  
determined isoform"
    match: proteins: Sw:P52736 Sw:Q06730 Sw:P51523 Sw:Q02386  
Sw:P51786 Sw:Q09676 Tr:Q06792 Sw:Q03923 Tr:Q0Y473  
Tr:Q0YK04 Tr:Q0P004 Sw:Q05201 Tr:Q0JL03 Sw:Q062396  
Tr:Q022X6 Tr:Q15322 Tr:Q0WV10 Tr:Q14913 Sw:Q01751  
Tr:Q0P0L1 Tr:Q0NS42 Sw:Q0YU03 Tr:P70591 Tr:Q14586  
Sw:Q02975 Tr:Q04892 Tr:Q0NS01 Tr:Q02977 Tr:Q0Y2N8  
Tr:Q0Y2N9 Tr:Q043296 Tr:Q0NMY4 Tr:Q0Y3W9 Tr:Q0NR11  
Sw:Q00765 Tr:Q05765 Tr:Q0QX19 Tr:Q0XSRI"
    /codon_start=1
    /evidence=not_experimental
    /product="dJ734P14.1 (novel KRAB box protein)"
    /protein_id="CA846713.2"
    /db_xref="GI:1160044"
    /translation="MMLPTPSALGDIYVEIILLPKNGEVEYTKKLTQNHKAKGPSN  
DIDCDQKKEGKAQIVPVRDVTIFTEAEKRRUSPEGRNLKYKEMLNRYNLLSLG  
QEETILANIVKSHLY"
    repeat_region 1153..1180
        /note="MER70B repeat: matches 267. .294 of consensus"
    repeat_region 2150..2632
        /note="L2 repeat: matches 2288. .2749 of consensus"
    repeat_region 2735..2975
        /note="AluSg repeat: matches 63. .303 of consensus"
    repeat_region 3026..3345
        /note="MER58B repeat: matches 2. .341 of consensus"
    repeat_region 3356..3684
        /note="MER33 repeat: matches -16. .324 of consensus"
    repeat_region 4121..4296
        /note="AluSg repeat: matches 135. .310 of consensus"
    repeat_region 4822..4952
        /note="FLAM_C repeat: matches 1. .133 of consensus"
    repeat_region 5086..5435
        /note="MLTII repeat: matches 28. .396 of consensus"
```


repeat__region 5775. 6061
/note="Aluub repeat: matches 1. .286 of consensus"
6324. 6691
/note="L1ME repeat: matches 4762. .5145 of consensus"
repeat__region 6727. 7278
/note="L1ME repeat: matches 5219. .5714 of consensus"
7592. 7742
/note="L1MC4 repeat: matches 6190. .6343 of consensus"
7743. 7941
/note="MER58A repeat: matches 4. .223 of consensus"
7942. 8042
/note="L1MC4 repeat: matches 6343. .6435 of consensus"
8043. 8330
/note="AluSp repeat: matches 1. .298 of consensus"
8331. 8396
/note="L1MC4 repeat: matches 6435. .6501 of consensus"
9563. 9650
/note="L2 repeat: matches 2611. .2698 of consensus"
<10150. 13007
/gene="dJ734P14.5"
/note="match: ESTs: Em:AV686588 Em:AV686506 Em:AV688199
Em:AV684693 Em:AV698692 Em:AV691116 Em:AV691591
Em:AV685994 Em:AM249885 Em:AI335243 Em:BE251006
Em:AV659318 Em:AI090631 Em:AI032856 Em:BE057295
Em:AI741853 Em:BE255957 Em:AI823988 Em:AM152481
Em:AM014719 Em:BE257108 Em:BE375372 Em:AV720356 Em:T06312
Em:BE513012 Em:AM299561 Em:AM250546 Em:AM592867
Em:BE250978 Em:W90340 Em:AI533990 Em:W90339 Em:BE744885
Em:AI042252 Em:BE250926 Em:AI719459"
/product="dJ734P14.5 (novel C2H2 type zinc finger protein)"
evidence=not_experimental
10150. 13007
/gene="dJ734P14.5"
<10150. 11663
/gene="dJ734P14.5"
/note="protein homologies suggest that this gene and gene
1. are part of the same gene in the form of a yet to be
determined isoform
some protein homologies suggest this could be a
pseudogene
match: protein: Tr:Q9Y3M9 Sw:007230 Sw:P52736 Tr:Q9Y3Y5
Sw:Q06730 Sw:Q99676 Tr:Q9Y611 Tr:Q14587 Tr:Q60792
Sw:PS1523 Tr:Q9NQY7 Sw:PS1814 Tr:Q9UL36 Tr:Q9UL37
Sw:Q31923 Tr:Q9Z2X6 Tr:Q9U113 Tr:Q62512 Tr:Q9UJW8
Sw:P52737 Tr:Q9NS43 Tr:Q9NQMS Tr:Q14585 Tr:Q75346
Sw:PS1786 Sw:Q16587 Tr:Q14709 Tr:Q94892 Tr:Q64247
Tr:Q9W747 Tr:Q02313 Tr:Q9NR04 Tr:Q9Y227 Tr:Q9P255
Tr:Q9Y603 Sw:PS15620 Tr:Q9XSR1"
/codon_start=3
evidence=not_experimental
/product="dJ734P14.5 (novel C2H2 type zinc finger protein)"
/protein_id="CAC32827.1"
/db_xref="GI:13160045"
/translation="NFHSSAEKPEIYTCSSCLAFSCQFLQSHVLDIFILCAENH
FHPGSSGGMKQOQOYSHVSCNFEAEGERGGSGPMSAREERTSAFSPDLQ
RPSADPGRNMVETEPSSAORPNVOLDKGLKLETPAFGINCEKEEPKHNESNN
ITNPITLLGKPKYICSDCGSRPKDRSTLIRHRIHSMKRPYVSCGCGEFSKSL
HORTHSEKPYLCRCGGSFNRKSLINHHQHTHSEKPYVSCGCGEFSKSLFRIH
RTHSEKPYVLCRCGGSFNRKSLIRKHQRIHSGEPYVSCGCGEFSKSLIKQRIH
HIDEPYVRCGGRGFCDSKTLIHRTSHGSKPYVCEGCGEFSKSLIVHORTHG
GKHVYVRCGGRGFSOKSLIHRTSHGSKPYVCEGCGEFSKSLIVHRTSHG
KRYVYVCEGGRGFSKSLIVHRTSHGSKPYVCEGCGEFSKSLINHORTH"
complement(10354. 10858)
/note="match: GSS: Em:AQ594268"
12315. 12460
/note="LTR16A repeat: matches 216. .353 of consensus"
complement(12523. 13002)
/note="match: STS: Em:AA463461"
12666. 13089
/note="match: GSS: Em:AQ080563"
12972. 12977
polYA_signal

	polyA_site	/gene="dJ734P14.5"	13007	/gene="dJ734P14.5"
	misc_feature	/note="match: GSS: Em:B32856"	13342..13761	
	repeat_region	/note="33 copies 2 mer tc 74% conserved"	13362..13427	
	repeat_region	/note="WSTD repeat: matches 1..394 of consensus"	14122..14511	
	repeat_region	/note="FLAM_C repeat: matches 2..132 of consensus"	15488..15621	
	misc_feature	/note="match: GSS: Em:AQ267892"	15912..15946	
	repeat_region	/note="Alusx repeat: matches 27..295 of consensus"	15987..16255	
	repeat_region	/note="FLAM_C repeat: matches 9..134 of consensus"	17043..17169	
	repeat_region	/note="UTRA1 repeat: matches 1..178 of consensus"	17809..17973	
	repeat_region		18009..18106	
	Query Match	1.0%; Score 28; DB 9; Length 90220;		
	Best Local Similarity	100.0%; Pred. No. 0.00044;		
	Matches	28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	1659 caccagaggacactcaaggagaagc	1686		
Dd	11310 CACCAGAGCACACTCAGGCGAGAAGC	11337		
RESULT 44	AF159227/C	109217 bp	DNA	linear PRI 05-FEB-2002Z
LOCUS	Homo sapiens chromosome 8 clone GSI-225N9	map 8q24,	complete	
ACCESSION	AF159227.6	GI:18497062		
VERSION	HTG; HTGS_FUILLTOP; HTGS_ACTIVEFIN.			
KEYWORDS	human.			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	AUTHORS	1 (bases 1 to 109217)		
TITLE	JOURNAL	Submitted (15-JUN-1999) Genome Analysis; Institute of Molecular Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany		
REFERENCE	AUTHORS	2 (bases 1 to 109217)		
TITLE	JOURNAL	Submitted (24-OCT-2001) Genome Analysis; Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany		
COMMENT		Center: Institute of Molecular Biotechnology		
		Web site: http://genome.jmb-jena.de/		
		Contact: gscf-submit@genome.jmb-jena.de		
		Project Information		
		Center project name: IMB		

Severy, P., Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Direct Submission
Submitted (31-MAY-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 143799)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Choepel, V., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Garmy, S., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., Lacroque, K., Maclean, C., Macdonald, P., Margulis, N., Matthews, C., McCarthy, M., McKernan, K., McPheters, R., Meldrum, J., Menus, L., Milnova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Direct Submission
Submitted (02-JUN-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
5 (bases 1 to 143799)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Alt, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, V., Colangelo, M., Collins, S., Collymore, A., Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Garmy, S., Glade, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Lacroque, K., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McKernan, K., McPheters, R., Meldrum, J., Menus, L., Milnova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (10-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 10, 2001 this sequence version replaced gi:14269782.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center project name: L3150
Center clone name: 11_F_23

Location/Qualifiers

source
1. 143799
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP11-11F23"
/clone.lib="RPCT-11 Human Male BAC"
744 . 925
repeat_region
1083 . 1140
/rpt_family="L3"
repeat_region
1586 . 1814
/rpt_family="AluSg"
repeat_region
1912 . 3816
/rpt_family="L1McC"
repeat_region
4170 . 4300
/rpt_family="L1McC"
repeat_region
5546 . 5853
/rpt_family="AluSg"
6814 . 7113
/rpt_family="AluY"
repeat_region
7304 . 7585
/rpt_family="L1PA7"
repeat_region
9689 . 9812
/rpt_family="L3"
9963 . 10092
/rpt_family="AluY"
repeat_region
10936 . 11160
/rpt_family="L3"
11196 . 11263
/rpt_family="MIR"
11292 . 11585
/rpt_family="AluB"
repeat_region
11795 . 12121
/rpt_family="AluSg"
repeat_region
12152 . 12483
/rpt_family="MER7A"
12923 . 13014
/rpt_family="MER31A"
12995 . 13140
/rpt_family="Tigger3(Colem)"
13146 . 13416
/rpt_family="MER31A"
repeat_region
13584 . 13922
/rpt_family="MER1B"
14939 . 14970
/rpt_family="(?)n"
repeat_region
16283 . 16380
/rpt_family="MIR3"
16440 . 16707
/rpt_family="AluX"
17310 . 17464
/rpt_family="MIR"
repeat_region
17753 . 17830
/rpt_family="MIR"
18533 . 18717
/rpt_family="MIR3"
18729 . 18916
/rpt_family="MIR3"
repeat_region
18923 . 19163
/rpt_family="L2"
repeat_region
20725 . 20764
/rpt_family="L2"
repeat_region
20850 . 20988
/rpt_family="EFAM.C"
repeat_region
21135 . 21763
/rpt_family="L2"
22218 . 22261
/rpt_family="CF-rich"
repeat_region
22264 . 22529
/rpt_family="AluX"

repeat_region 23242..23396
 /rpt_family="L2"
 repeat_region 23628..23700
 /rpt_family="L2"
 repeat_region 25505..25550
 /rpt_family="(TC)n"

Query Match 1.0%; Score 28; DB 9; Length 143799;
 Best Local Similarity 100.0%; Pred. No. 0.00044;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2358 tgaatgagcagagagcgaatgc 2385
 ||||||||||||||||||||||||||||
 Db 123323 TGAATGAGCAGAGCAGCAATGCC 123350

RESULT 47
 AC010998 146831 bp DNA linear PRI 09-AUG-2001
 DEFINITION Homo sapiens chromosome 10 clone RP11-95116, complete sequence.
 AC010998
 VERSION AC010998.13 GI:15144282
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 146831)
 Smith,D.R.

TITLE Genome Therapeutics Corporation Sequencing Center: Human Genome
 Sequence Data

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 146831)
 Smith,D.R.

TITLE Direct Submission
 JOURNAL Submitted (29-SEP-1999) Genome Therapeutics Corporation, 100 Beaver
 Street, Waltham, MA 02453, USA

REFERENCE 3 (bases 1 to 146831)
 Smith,D.R.

TITLE Direct Submission
 JOURNAL Submitted (09-AUG-2001) Genome Therapeutics Corporation, 100 Beaver
 Street, Waltham, MA 02453, USA

COMMENT On Aug 9, 2001 this sequence version replaced gi:14717312.
 FEATURES
 source

1..146831
 Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="10"

/clone="RP11-95116"

/clone_1b="RP11-95116"

BASE COUNT 41832 a 31937 c 31285 g 41777 t

ORIGIN

Query Match 1.0%; Score 28; DB 9; Length 146831;
 Best Local Similarity 100.0%; Pred. No. 0.00044;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2511 acccttacctatttaccgatactgc 2538
 ||||||||||||||||||||||||||||

Db 119994 ACCCTTACCTATTATTACGTATCGTC 120021

RESULT 48

AC021765/c

LOCUS AC021765 152393 bp DNA linear HTG 20-SEP-2000

DEFINITION Homo sapiens clone RP11-9D19, WORKING DRAFT SEQUENCE, 11 unordered
 pieces.

AC021765

AC021765.3 GI:10198414

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 152393)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens, clone RP11-9D19
 Unpublished
 REFERENCE 2 (bases 1 to 152393)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
 Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
 Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
 Dearriano,K., Dewar,K., Domino,M., Doyle,M., Fenesor,J.,
 Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
 Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,D.,
 Landers,T., Lehoczký,J., Levine,R., Lien,C., Liu,G., Locke,K.,
 Macdonald,P., Marcuzzi,N., McEwan,P., McGurk,A., McInerlan,K.,
 McInerlan,R., Meldrum,D., Menus,L., Morrow,J., Naylor,J.,
 Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
 Pierre,N., Pisanil,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
 Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
 Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
 Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
 Zimmer,A. and Zody,M.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

NOTE: This is a 'working draft' sequence. It currently
 consists of 11 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.
 * 1 28910: contig of 28910 bp in length
 * 28911 29010: gap of 100 bp
 * 29011 31437: contig of 2427 bp in length
 * 31438 31537: gap of 100 bp
 * 31538 33952: contig of 2415 bp in length
 * 33953 34052: gap of 100 bp
 * 34053 38372: contig of 4320 bp in length
 * 38373 38472: gap of 100 bp
 * 38473 81141: contig of 42669 bp in length
 * 81142 81241: gap of 100 bp
 * 81242 86894: contig of 5553 bp in length
 * 86895 86994: gap of 100 bp
 * 86995 95082: contig of 8088 bp in length
 * 95083 95182: gap of 100 bp

Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L3002
 Center clone name: 9_D_19
 Summary Statistics
 Sequencing vector: M13: M7815: 100% of reads
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap: version 0.960731
 Consensus quality: 146933 bases at least Q40
 Consensus quality: 149520 bases at least Q30
 Consensus quality: 150601 bases at least Q20
 Insert size: 157000; agarose-fp
 Insert size: 151393; sum-of-contigs
 Quality coverage: 4.6 in Q20 bases; agarose-fp
 Quality coverage: 4.7 in Q20 bases; sum-of-contigs

* 95183 110046: contig of 14864 bp in length
 * 110047 110146: gap of 100 bp
 * 110147 129817: contig of 19671 bp in length
 * 129818 129917: gap of 100 bp
 * 129918 148772: contig of 18855 bp in length
 * 148773 148872: gap of 100 bp
 * 148873 152393: contig of 3521 bp in length.
 Location/Qualifiers
 1. 152393

FEATURES

source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="RP11-9D19"
 /clone.lib="RP11-11 Human Male BAC"

misc_feature

1. 28910
 /note="assembly-fragment
 clone_end:SP6
 vector_side:left"

misc_feature

29011..31437
 /note="assembly-fragment"

misc_feature

31538..33952
 /note="assembly-fragment"

misc_feature

34053..38372
 /note="assembly-fragment"

misc_feature

38473..81141
 /note="assembly-fragment"

misc_feature

81242..86894
 /note="assembly-fragment"

misc_feature

86995..95082
 /note="assembly-fragment"

misc_feature

95183..110046
 /note="assembly-fragment"

misc_feature

110147..129817
 /note="assembly-fragment"

misc_feature

129918..148772
 /note="assembly-fragment"

misc_feature

148873..152393
 /note="assembly-fragment"

misc_feature

vector_end:T7
 vector_side:right"

misc_feature

49730 a 27512 c 26497 g 47643 t 1011 others

BASE COUNT

49730 a 27512 c 26497 g 47643 t 1011 others

ORIGIN

Query Match

1.0%; Score 28; DB 2; Length 152393;

Best Local Similarity

100.0%; Pred. No. 0.00044;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2358 tgatagtgcagagagcagcgaatgcc 2385

Db 27507 TGATAGTGCAGAGCAGCAGCAATGCC 27480

RESULT 49

AC104974/c 154604 bp DNA linear HTG 13-FEB-2002
 LOCUS Homo sapiens chromosome 17 clone RP11-145L16 map 17, WORKING DRAFT
 DEFINITION SEQUENCE: 23 ordered pieces.
 AC104974
 AC104974.3 GI:18653683
 VERSION HTG: HTGS_PHASE2; HTGS_DRAFT; HTGS_FUZZTOP.
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 154604)
 Homo sapiens chromosome 17, clone RP11-145L16
 Unpublished
 2 (bases 1 to 154604)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
 Brown,A., Camarata,J., Campopiano,A., Chang,Y., Chazaro,B.,
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,

TITLE
JOURNAL

COMMENT

Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hages,B., Hearford,A., Horton,L., Hulme,W., Illiev,I., Johnson,R.,
 Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
 Maclean,C., MacDonald,P., Major,J., Margulis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrum,J.,
 Menus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Rella,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Roselli,M., Roy,A., Santos,R., Schauer,S., Schupbach,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,R., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Testfay,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zelnou,R., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (22-DEC-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Feb 13, 2002 this sequence version replaced gi:18464184.
 All repeats were identified using RepeatMasker:
 Smit,A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L22871
 Center clone name: 145_L16
 ----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 148031 bases at least Q40
 Consensus quality: 150405 bases at least Q30
 Consensus quality: 151462 bases at least Q20
 Insert size: 146000; agarose-fp
 Insert size: 152404; sum-of-ctdigs
 Quality coverage: 9.3 in Q20 bases; agarose-fp
 Quality coverage: 8.9 in Q20 bases; sum-of-ctdigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 23 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submittor.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 1709: contig of 1709 bp in length
 * 1710 1809: gap of 100 bp
 * 1810 1999: contig of 190 bp in length
 * 2000 2099: gap of 100 bp
 * 2100 3127: contig of 1028 bp in length
 * 3128 3227: gap of 100 bp
 * 3228 4561: contig of 1334 bp in length
 * 4562 4661: gap of 100 bp
 * 4662 5811: contig of 1150 bp in length
 * 5812 5911: gap of 100 bp
 * 5912 7151: contig of 1240 bp in length
 * 7152 7251: gap of 100 bp
 * 7252 8261: contig of 1010 bp in length
 * 8262 8361: gap of 100 bp
 * 8362 9739: contig of 1378 bp in length
 * 9740 9839: gap of 100 bp
 * 9840 12105: contig of 2265 bp in length
 * 12106 12205: gap of 100 bp
 * 12206 35760: contig of 23555 bp in length

```
* 35761 35860: gap of 100 bp
* 35861 38000: contig of 2140 bp in length
* 38001 38100: gap of 100 bp
* 38101 41369: contig of 3269 bp in length
* 41370 41469: gap of 100 bp
* 41470 44094: contig of 2625 bp in length
* 44095 44194: gap of 100 bp
* 44195 47400: contig of 3206 bp in length
* 47401 47500: gap of 100 bp
* 47501 53741: contig of 6241 bp in length
* 53742 53841: gap of 100 bp
* 53842 60833: contig of 6992 bp in length
* 60834 60933: gap of 100 bp
* 60934 72421: contig of 11488 bp in length
* 72422 72521: gap of 100 bp
* 72522 84887: contig of 12366 bp in length
* 84888 84987: gap of 100 bp
* 84988 99600: contig of 14613 bp in length
* 99601 99700: gap of 100 bp
* 99701 115353: contig of 15653 bp in length
* 115354 115453: gap of 100 bp
* 115454 131255: contig of 15802 bp in length
* 131256 131355: gap of 100 bp
* 131356 152505: contig of 21150 bp in length
* 152506 152605: gap of 100 bp
* 152606 154604: contig of 1999 bp in length.
```

FEATURES

source

```
1. 154604
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone_lib="RP11-145L16"
/clone_id="RP11-11 Human Male BAC"
1. 1709
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
1810. 1999
/note="assembly_fragment"
2100. 3127
/note="assembly_fragment"
3228. 4561
/note="assembly_fragment"
4662. 5811
/note="assembly_fragment"
5912. 7151
/note="assembly_fragment"
7252. 8261
/note="assembly_fragment"
8362. 9739
/note="assembly_fragment"
9840. 12105
/note="assembly_fragment"
12206. 35760
/note="assembly_fragment"
35861. 38000
/note="assembly_fragment"
38101. 41369
/note="assembly_fragment"
41470. 44094
/note="assembly_fragment"
44195. 47400
/note="assembly_fragment"
47501. 53741
/note="assembly_fragment"
53842. 60833
/note="assembly_fragment"
60934. 72421
/note="assembly_fragment"
72522. 84887
/note="assembly_fragment"
84988. 99600
misc_feature
```

```
/note="assembly_fragment"
99701. 115353
/note="assembly_fragment"
115454. 131255
/note="assembly_fragment"
131356. 152505
/note="assembly_fragment"
152606. 154604
/note="assembly_fragment"
clone_end:T7
vector_side:right"
BASE COUNT 42238 a 34832 c 34671 g 40661 t 2202 others
ORIGIN
```

```
Query Match 1.0%: Score 28; DB 2; Length 154604;
Best Local Similarity 100.0%; Pred No. 0.00044;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2461 aattgaacatgcttcctccattggtg 2488
|||||
Db 28516 AATTGAGAACCTGCTTCCCATTTGGTG 28489
```

```
RESULT 50
AC080121/c 154987 bp DNA linear PRI 27-DBC-2001
LOCUS Homo sapiens chromosome 15, clone RP13-500A21, complete sequence.
AC080121
ACCESSION AC080121
VERSION AC080121.4 GI:17646926
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
```

```
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 154987) Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE 1
JOURNAL Birren,B., Linton,L., Nusbaum,C. and Lander,E.
REFERENCE Homo sapiens chromosome 15, clone RP13-500A21
AUTHORS Unpublished
2 (bases 1 to 154987)
```

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavsky,L., Bouknight,B., Brown,A., Burkett,G., Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearliano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Hagsos,B., Heatford,A., Horton,L., Iliev,I., Johnson,R., Jones,C., Kamm,L., Karatas,A., Lacroque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K., MCPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivat,T.M., Oliver,J., Peterson,K., Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Riebeck,M., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Sougez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submision
Submitted (24-SEP-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 154987)

REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouknight,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearliano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagsos,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K.,

Lamazares, R., Landers, T., Lehocsky, J., Levine, R., Liu, G., Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Relta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wymann, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (27-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 13, 2001 this sequence version replaced gi:17386327.
All repeats were identified using RepeatMasker:
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L11140
Center clone name: 500_A_21

FEATURES
source

Location/Qualifiers
1.154967
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="rp13-500A21"
/clone_lib="RPCT-13 Human Female BAC 481-768"
2.315
/rpt_family="MSTB1"
497..527
/note="<30 qual SNGL region"
551..559
/note="<30 qual SNGL region"
596..607
/note="<30 qual SNGL region"
632..635
/note="<30 qual SNGL region"
636..642
/note="Single clone coverage"
643..665
/note="<30 qual single clone coverage"
666..671
/note="Single clone coverage"
672..679
/note="Single clone coverage"
680..686
/note="<30 qual single clone coverage"
687..719
/note="Single clone coverage"
720..725
/note="<30 qual single clone coverage"
726..761
/note="Single clone coverage"
765..770
/note="<30 qual SNGL region"
790..794
/note="<30 qual SNGL region"
801..818
/note="<30 qual SNGL region"
837..842
/note="<30 qual SNGL region"
1821..1995
/rpt_family="CT-rich"
repeat_region

repeat_region 1987..2054
/rpt_family="CT-rich"
repeat_region 2050..2130
/rpt_family="(TTTC)n"
2076..2080
/note="<30 qual SNGL region"
unsure
repeat_region
/rpt_family="MIR"
2350..2539
3465..3517
/rpt_family="MIR"
repeat_region
/rpt_family="MIR"
3606..3922
/rpt_family="AluJo"
repeat_region
/rpt_family="L1ME3"
4820
/rpt_family="MERS5B"
4830..5433
/rpt_family="L1MC4"
5433..5596
/rpt_family="L1MC4a"
6265..6666
/rpt_family="L1MDa"
6802..6916
/rpt_family="L1MC4a"
7884..7977
/rpt_family="MERS3"
8199..8435
/rpt_family="L1R16A1"
8463..8704
/rpt_family="Aluub"
9157..10047
/rpt_family="L1PA9"
10048..10395
/rpt_family="TME1B"
10396..15394
/rpt_family="L1PA9"
16029..16078
/rpt_family="(TA)n"
16089..16209
/rpt_family="MIR"
17153..17203
/rpt_family="CT-rich"
17213..17572
/rpt_family="MTR1A1"
18363..18511
/rpt_family="MIR"
19022..19315
/rpt_family="AluY"
19886..20138
/rpt_family="L2"
20497
/rpt_family="AluSg"
21560..21629
/rpt_family="L2"
21682..21991
/rpt_family="AluSg"
22192..22508
/rpt_family="MTR1A1"
23171..23589
/rpt_family="MSTB"
24349..24415
/rpt_family="L1MC4a"
24416..24482
/rpt_family="(TG)n"
24483..24861
/rpt_family="L1MC4a"
24868..24891
/rpt_family="AT-rich"
24892..25021
/rpt_family="Aluub"
25022..25376
repeat_region

```

repeat_region      /rpt_family="L1MC4a"
                    25475..25945
repeat_region      /rpt_family="MLT1G1"
                    complement(25963..26383)
repeat_region      /rpt_family="L1MB3"
                    26475..26635
                    /rpt_family="L1MC4a"
    
```

```

Query Match      1.0%; Score 28; DB 9; Length 154987;
Best Local Similarity 100.0%; Pred. No. 0.00044;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    
```

```

QY 2358 tgatagtgacagggcgagtcgaatgcc 2385
      ||||||||||||||||||||||||||||
Db 88161 TGATAGTGCAGCAGCGAGTCGAATGCC 88134
    
```

Search completed: May 16, 2002, 11:48:37
 Job time: 12472 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2002, 08:22:45 ; Search time 93.58 Seconds
(without alignments)
7276.085 Million cell updates/sec

Title: US-09-898-556A-3
Perfect score: 2772
Sequence: 1 cagcgcgcttaagctggtg.....tttaccacacctaccct 2772

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/lna/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/lna/Dackfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	0.8	2382	4	US-09-588-256-9
2	20	0.7	1892	2	US-08-933-750C-66
3	20	0.7	1892	3	US-09-234-613-66
4	19	0.7	3240	4	US-09-262-773-7
5	19	0.7	3244	4	US-09-262-773-5
6	19	0.7	3264	4	US-09-262-773-5
7	19	0.7	3268	4	US-09-262-773-1
8	19	0.7	3810	2	US-08-475-844-8
9	19	0.7	3810	5	PCT-US95-08429-8
10	19	0.7	20137	4	US-09-262-773-206
11	19	0.7	20138	4	US-09-262-773-9
12	19	0.7	23071	4	US-08-724-394A-20
13	19	0.7	246240	2	US-08-724-394A-21
14	19	0.7	246240	2	US-08-724-394A-22
15	19	0.7	246240	2	US-08-724-394A-22
16	18	0.6	575	1	US-08-507-016-8
17	18	0.6	1027	4	US-09-303-524A-1
18	18	0.6	1558	1	US-08-416-870C-9
19	18	0.6	2133	2	US-08-820-170A-11
20	18	0.6	2133	3	US-09-055-699-11
21	18	0.6	2133	4	US-09-273-565-11
22	18	0.6	2133	4	US-09-565-538-11
23	18	0.6	3754	2	US-08-820-170A-12
24	18	0.6	3754	3	US-09-055-699-12
25	18	0.6	3754	4	US-09-273-565-12
26	18	0.6	3754	4	US-09-565-538-12
27	18	0.6	7886	2	US-08-751-189-2

c 28	18	0.6	7886	2	US-09-060-836-2	Sequence 2, Appl
c 29	18	0.6	7886	4	US-09-184-445-2	Sequence 2, Appl
c 30	17	0.6	80	5	PCT-US91-02942-99	Sequence 99, Appl
c 31	17	0.6	273	3	US-08-956-182-21	Sequence 21, Appl
c 32	17	0.6	289	1	US-08-592-126-124	Sequence 124, Appl
c 33	17	0.6	347	4	US-08-905-223-133	Sequence 133, Appl
c 34	17	0.6	532	4	US-09-328-111-648	Sequence 648, Appl
c 35	17	0.6	563	4	US-09-385-982-433	Sequence 433, Appl
c 36	17	0.6	732	3	US-08-956-182-16	Sequence 16, Appl
c 37	17	0.6	1362	4	US-09-171-461-47	Sequence 47, Appl
c 38	17	0.6	1561	1	US-07-662-007B-38	Sequence 38, Appl
c 39	17	0.6	1561	1	US-07-968-971A-1	Sequence 1, Appl
c 40	17	0.6	1561	1	US-07-824-247-38	Sequence 38, Appl
c 41	17	0.6	1561	1	US-07-824-247-41	Sequence 41, Appl
c 42	17	0.6	1561	1	US-08-142-474A-1	Sequence 1, Appl
c 43	17	0.6	1561	1	US-08-424-406-1	Sequence 1, Appl
c 44	17	0.6	1561	1	US-08-464-523B-5	Sequence 5, Appl
c 45	17	0.6	1561	1	US-08-469-203A-1	Sequence 1, Appl
c 46	17	0.6	1561	1	US-08-469-203A-1	Sequence 1, Appl
c 47	17	0.6	1561	3	US-08-470-204A-38	Sequence 38, Appl
c 48	17	0.6	1561	3	US-08-470-204A-41	Sequence 41, Appl
c 49	17	0.6	2031	2	US-08-933-750C-53	Sequence 53, Appl
c 50	17	0.6	2031	3	US-09-234-613-53	Sequence 53, Appl
c 51	17	0.6	2680	4	US-09-063-035-1	Sequence 1, Appl
c 52	17	0.6	2971	1	US-07-718-575-7	Sequence 7, Appl
c 53	17	0.6	2971	1	US-08-481-206-7	Sequence 7, Appl
c 54	17	0.6	5022	2	US-08-486-269A-7	Sequence 7, Appl
c 55	17	0.6	5022	2	US-07-289-368-1	Sequence 1, Appl
c 56	17	0.6	6953	1	US-07-805-123C-2	Sequence 2, Appl
c 57	17	0.6	6953	1	US-08-033-081B-2	Sequence 2, Appl
c 58	17	0.6	7333	4	US-08-766-528-2	Sequence 2, Appl
c 59	17	0.6	43804	4	US-09-171-461-1	Sequence 1, Appl
c 60	17	0.6	49136	4	US-09-422-869-1	Sequence 1, Appl
c 61	17	0.6	72928	3	US-09-009-913-1	Sequence 1, Appl
c 62	17	0.6	90050	4	US-09-245-041-5	Sequence 5, Appl
c 63	16	0.6	18	3	US-09-156-807-40	Sequence 40, Appl
c 64	16	0.6	28	2	US-09-289-368-2	Sequence 2, Appl
c 65	16	0.6	31	1	US-08-510-215A-11	Sequence 11, Appl
c 66	16	0.6	41	3	US-08-783-853A-101	Sequence 101, Appl
c 67	16	0.6	54	4	US-08-584-040-8384	Sequence 8284, Appl
c 68	16	0.6	256	1	US-08-616-368A-25	Sequence 25, Appl
c 69	16	0.6	256	3	US-09-054-298-25	Sequence 25, Appl
c 70	16	0.6	256	4	US-08-818-655-25	Sequence 25, Appl
c 71	16	0.6	285	1	US-08-469-421-6	Sequence 6, Appl
c 72	16	0.6	285	1	US-08-250-975-6	Sequence 6, Appl
c 73	16	0.6	285	2	US-08-605-002A-6	Sequence 6, Appl
c 74	16	0.6	285	2	US-08-950-449A-6	Sequence 6, Appl
c 75	16	0.6	345	5	PCT-US94-10529-6	Sequence 6, Appl
c 76	16	0.6	360	1	US-08-592-126-113	Sequence 113, Appl
c 77	16	0.6	360	1	US-08-318-970B-48	Sequence 48, Appl
c 78	16	0.6	396	1	US-08-181-271A-102	Sequence 102, Appl
c 79	16	0.6	396	1	US-08-449-315-102	Sequence 102, Appl
c 80	16	0.6	396	1	US-08-449-803-102	Sequence 102, Appl
c 81	16	0.6	396	1	US-08-449-043-102	Sequence 102, Appl
c 82	16	0.6	396	1	US-08-456-265A-102	Sequence 102, Appl
c 83	16	0.6	396	1	US-08-455-416-102	Sequence 102, Appl
c 84	16	0.6	396	1	US-08-455-244-102	Sequence 102, Appl
c 85	16	0.6	396	1	US-08-454-876-102	Sequence 102, Appl
c 86	16	0.6	396	2	US-08-457-364-102	Sequence 102, Appl
c 87	16	0.6	396	2	US-08-456-262-102	Sequence 102, Appl
c 88	16	0.6	396	2	US-08-456-240-102	Sequence 102, Appl
c 89	16	0.6	396	2	US-08-455-736-102	Sequence 102, Appl
c 90	16	0.6	396	2	US-08-971-217-102	Sequence 102, Appl
c 91	16	0.6	396	4	US-09-350-600-102	Sequence 102, Appl
c 92	16	0.6	399	3	US-09-188-930-63	Sequence 63, Appl
c 93	16	0.6	498	1	US-07-781-255A-14	Sequence 14, Appl
c 94	16	0.6	542	4	US-09-305-635-5	Sequence 5, Appl
c 95	16	0.6	566	4	US-09-385-982-388	Sequence 388, Appl
c 96	16	0.6	620	4	US-08-998-416-42	Sequence 42, Appl
c 97	16	0.6	699	4	US-08-998-416-591	Sequence 591, Appl
c 98	16	0.6	709	4	US-08-998-416-599	Sequence 599, Appl
c 99	16	0.6	836	4	US-08-971-090-6	Sequence 6, Appl
c 100	16	0.6	836	4	US-08-971-090-7	Sequence 7, Appl

101	16	0.6	862	3	US-08-713-569-4	Sequence 4, Appl1	174	16	0.6	5319	1	US-08-169-927-1	Sequence 1, Appl1
102	16	0.6	1023	3	US-08-554-385-4	Sequence 4, Appl1	175	16	0.6	5816	4	US-09-220-641-4	Sequence 4, Appl1
103	16	0.6	1058	2	US-09-156-807-1	Sequence 1, Appl1	176	16	0.6	6645	2	US-08-380-403A-4	Sequence 4, Appl1
104	16	0.6	1100	2	US-08-950-449A-18	Sequence 18, Appl1	177	16	0.6	6645	2	US-08-895-628-4	Sequence 4, Appl1
105	16	0.6	1119	1	US-08-416-756A-1	Sequence 1, Appl1	178	16	0.6	7622	2	US-09-305-633-1	Sequence 1, Appl1
106	16	0.6	1200	1	US-08-011-398B-3	Sequence 3, Appl1	179	16	0.6	13146	2	US-08-724-354D-3	Sequence 3, Appl1
107	16	0.6	1200	1	US-08-464-051-3	Sequence 3, Appl1	180	16	0.6	13146	3	US-09-470-984A-3	Sequence 3, Appl1
108	16	0.6	1200	2	US-08-462-498-3	Sequence 3, Appl1	181	16	0.6	35100	1	US-08-306-691B-19	Sequence 19, Appl1
109	16	0.6	1200	3	US-08-617-860B-4	Sequence 4, Appl1	182	16	0.6	35100	2	PCT-US93-06251-19	Sequence 19, Appl1
110	16	0.6	1247	2	US-08-773-870-2	Sequence 2, Appl1	183	16	0.6	56516	2	US-08-996-306-1	Sequence 1, Appl1
111	16	0.6	1255	4	US-09-412-554A-7	Sequence 7, Appl1	184	16	0.6	56516	4	US-09-338-907-1	Sequence 1, Appl1
112	16	0.6	1309	2	US-08-933-750C-63	Sequence 63, Appl1	185	16	0.6	56516	4	US-09-218-207-1	Sequence 1, Appl1
113	16	0.6	1309	3	US-09-334-613-63	Sequence 63, Appl1	186	16	0.6	56520	4	US-09-338-907-1	Sequence 1, Appl1
114	16	0.6	1338	1	US-07-781-254A-6	Sequence 6, Appl1	187	16	0.6	56520	4	US-09-218-207-1	Sequence 1, Appl1
115	16	0.6	1338	1	US-07-781-254A-7	Sequence 7, Appl1	188	16	0.6	71989	4	US-09-443-501A-2	Sequence 2, Appl1
116	16	0.6	1381	2	US-08-950-449A-13	Sequence 13, Appl1	189	16	0.6	152331	3	US-09-128-155-16	Sequence 16, Appl1
117	16	0.6	1423	1	US-08-469-421-13	Sequence 13, Appl1	190	16	0.6	176373	3	US-09-128-155-17	Sequence 17, Appl1
118	16	0.6	1423	1	US-08-250-975-13	Sequence 13, Appl1	191	16	0.6	4403765	4	US-09-103-840A-2	Sequence 2, Appl1
119	16	0.6	1423	2	US-08-605-002A-13	Sequence 13, Appl1	192	16	0.6	4411529	4	US-09-103-840A-1	Sequence 1, Appl1
120	16	0.6	1423	5	PCT-US94-10529-13	Sequence 13, Appl1	193	15	0.5	22	4	US-09-240-918-24	Sequence 24, Appl1
121	16	0.6	1470	4	US-09-099-041A-27	Sequence 27, Appl1	194	15	0.5	23	1	US-08-471-426-13	Sequence 13, Appl1
122	16	0.6	1516	3	US-09-248-528-3	Sequence 3, Appl1	195	15	0.5	23	1	US-08-471-426-13	Sequence 13, Appl1
123	16	0.6	1516	4	US-09-549-108-3	Sequence 3, Appl1	196	15	0.5	23	3	US-08-872-094-7	Sequence 7, Appl1
124	16	0.6	1516	4	US-09-549-111-3	Sequence 3, Appl1	197	15	0.5	23	5	PCT-US94-01709-13	Sequence 13, Appl1
125	16	0.6	1516	4	US-09-549-106-3	Sequence 3, Appl1	198	15	0.5	24	1	US-08-211-205-5	Sequence 5, Appl1
126	16	0.6	1516	4	US-09-550-394-3	Sequence 3, Appl1	199	15	0.5	24	2	US-08-273-146-11	Sequence 11, Appl1
127	16	0.6	1516	4	US-09-426-868-1	Sequence 1, Appl1	200	15	0.5	30	1	US-08-478-039-47	Sequence 47, Appl1
128	16	0.6	1517	3	US-09-248-528-2	Sequence 2, Appl1	201	15	0.5	30	1	US-08-149-099C-5	Sequence 5, Appl1
129	16	0.6	1517	4	US-09-549-108-2	Sequence 2, Appl1	202	15	0.5	30	1	US-08-476-344A-47	Sequence 47, Appl1
130	16	0.6	1517	4	US-09-549-111-2	Sequence 2, Appl1	203	15	0.5	30	1	US-08-476-273-9	Sequence 9, Appl1
131	16	0.6	1517	4	US-09-549-106-2	Sequence 2, Appl1	204	15	0.5	30	2	US-08-478-966A-5	Sequence 5, Appl1
132	16	0.6	1517	4	US-09-550-394-2	Sequence 2, Appl1	205	15	0.5	30	3	US-08-523-894-40	Sequence 40, Appl1
133	16	0.6	1539	3	US-09-286-904-1	Sequence 1, Appl1	206	15	0.5	35	2	US-08-561-521-19	Sequence 19, Appl1
134	16	0.6	1636	4	US-09-039-198A-1	Sequence 1, Appl1	207	15	0.5	35	3	US-09-070-637-18	Sequence 18, Appl1
135	16	0.6	1643	2	US-08-486-839-3	Sequence 3, Appl1	208	15	0.5	35	5	PCT-US95-01219-19	Sequence 19, Appl1
136	16	0.6	1643	3	US-09-151-011-3	Sequence 3, Appl1	209	15	0.5	36	1	US-08-478-039A-46	Sequence 46, Appl1
137	16	0.6	1643	4	US-09-343-623-3	Sequence 3, Appl1	210	15	0.5	36	1	US-08-478-342A-46	Sequence 46, Appl1
138	16	0.6	1656	4	US-09-039-198A-3	Sequence 3, Appl1	211	15	0.5	36	3	US-08-523-894-39	Sequence 39, Appl1
139	16	0.6	1656	4	US-09-247-155-72	Sequence 72, Appl1	212	15	0.5	39	1	US-08-236-520-5	Sequence 5, Appl1
140	16	0.6	1686	4	US-08-029-170-30	Sequence 30, Appl1	213	15	0.5	39	1	US-07-977-696C-9	Sequence 9, Appl1
141	16	0.6	1713	2	US-08-486-839-5	Sequence 5, Appl1	214	15	0.5	39	1	US-08-129-930B-9	Sequence 9, Appl1
142	16	0.6	1713	3	US-09-151-011-5	Sequence 5, Appl1	215	15	0.5	39	4	US-08-976-288A-9	Sequence 9, Appl1
143	16	0.6	1713	4	US-09-343-623-5	Sequence 5, Appl1	216	15	0.5	39	5	PCT-US95-05262-5	Sequence 5, Appl1
144	16	0.6	1780	2	US-08-933-821-5	Sequence 5, Appl1	217	15	0.5	42	1	US-08-375-116A-39	Sequence 39, Appl1
145	16	0.6	1780	4	US-09-136-828-5	Sequence 5, Appl1	218	15	0.5	42	2	US-08-273-146-25	Sequence 25, Appl1
146	16	0.6	1863	1	US-07-781-254A-4	Sequence 4, Appl1	219	15	0.5	47	2	US-08-659-567-7	Sequence 7, Appl1
147	16	0.6	1863	1	US-07-864-475A-5	Sequence 5, Appl1	220	15	0.5	48	1	US-08-017-570-16	Sequence 16, Appl1
148	16	0.6	2051	1	US-07-864-475A-3	Sequence 3, Appl1	221	15	0.5	48	1	US-08-211-202-7	Sequence 7, Appl1
149	16	0.6	2051	2	US-08-468-249A-3	Sequence 3, Appl1	222	15	0.5	48	1	US-08-471-426-16	Sequence 16, Appl1
150	16	0.6	2168	2	US-08-633-879C-1	Sequence 1, Appl1	223	15	0.5	48	2	US-08-800-199-12	Sequence 12, Appl1
151	16	0.6	2222	1	US-07-828-790B-5	Sequence 5, Appl1	224	15	0.5	48	3	US-09-296-595-12	Sequence 12, Appl1
152	16	0.6	2308	2	US-08-480-229C-28	Sequence 28, Appl1	225	15	0.5	48	5	PCT-US94-011709-16	Sequence 16, Appl1
153	16	0.6	2308	2	US-08-659-235C-28	Sequence 28, Appl1	226	15	0.5	51	3	US-08-589-939-60	Sequence 60, Appl1
154	16	0.6	2698	2	US-08-677-298-1	Sequence 28, Appl1	227	15	0.5	51	3	US-09-046-247-27	Sequence 27, Appl1
155	16	0.6	2793	1	US-07-646-537B-1	Sequence 1, Appl1	228	15	0.5	54	2	US-08-448-418-70	Sequence 70, Appl1
156	16	0.6	2859	4	US-09-099-041A-9	Sequence 9, Appl1	229	15	0.5	57	1	US-08-057-937-5	Sequence 5, Appl1
157	16	0.6	2906	1	US-08-554-612C-49	Sequence 49, Appl1	230	15	0.5	57	4	US-09-057-937-5	Sequence 5, Appl1
158	16	0.6	2921	3	US-08-904-032-2	Sequence 2, Appl1	231	15	0.5	61	1	US-08-408-133-8	Sequence 8, Appl1
159	16	0.6	3080	4	US-09-099-041A-25	Sequence 25, Appl1	232	15	0.5	61	1	US-08-454-683-8	Sequence 8, Appl1
160	16	0.6	3114	4	US-08-946-026-12	Sequence 12, Appl1	233	15	0.5	61	2	US-08-454-680-8	Sequence 8, Appl1
161	16	0.6	3150	4	US-08-943-768-1	Sequence 1, Appl1	234	15	0.5	67	1	US-08-472-281A-42	Sequence 42, Appl1
162	16	0.6	3192	4	US-09-412-554A-1	Sequence 1, Appl1	235	15	0.5	67	1	US-08-472-281A-42	Sequence 42, Appl1
163	16	0.6	3382	4	US-09-099-041A-7	Sequence 7, Appl1	236	15	0.5	67	2	US-08-477-989B-42	Sequence 42, Appl1
164	16	0.6	3774	2	US-08-950-449A-11	Sequence 11, Appl1	237	15	0.5	81	3	US-08-933-983-55	Sequence 55, Appl1
165	16	0.6	3813	1	US-08-469-421-11	Sequence 11, Appl1	238	15	0.5	84	2	US-08-483-528B-69	Sequence 89, Appl1
166	16	0.6	3813	2	US-08-250-975-11	Sequence 11, Appl1	239	15	0.5	84	3	US-08-483-528B-69	Sequence 89, Appl1
167	16	0.6	3813	2	US-08-605-002A-11	Sequence 11, Appl1	240	15	0.5	90	6	5188961-7	Patent No. 5188961
168	16	0.6	4403	5	PCT-US94-10529-11	Sequence 11, Appl1	241	15	0.5	96	3	US-08-523-894-58	Sequence 58, Appl1
169	16	0.6	4403	2	US-08-284-941-1	Sequence 1, Appl1	242	15	0.5	97	1	US-08-377-495-4	Sequence 4, Appl1
170	16	0.6	4403	2	US-08-447-642-1	Sequence 1, Appl1	243	15	0.5	97	2	US-08-748-415-4	Sequence 4, Appl1
171	16	0.6	4403	2	US-09-336-503-1	Sequence 1, Appl1	244	15	0.5	97	3	US-08-747-654-4	Sequence 4, Appl1
172	16	0.6	4403	5	PCT-US93-02147A-1	Sequence 1, Appl1	245	15	0.5	97	4	US-09-235-353-4	Sequence 4, Appl1
173	16	0.6	4403	5	PCT-US93-02147A-1	Sequence 1, Appl1	246	15	0.5	98	3	US-08-479-744A-43	Sequence 43, Appl1

247	15	0.5	98	3	US-08-280-757B-43	Sequence 43, Appl	320	15	0.5	406	5	PCT-US95-01219-14	Sequence 14, Appl
c 248	15	0.5	99	3	US-08-479-744A-42	Sequence 42, Appl	321	15	0.5	408	1	US-08-408-133-5	Sequence 5, Appl
c 249	15	0.5	99	3	US-08-280-757B-42	Sequence 42, Appl	322	15	0.5	408	1	US-08-454-683-5	Sequence 5, Appl
c 250	15	0.5	118	2	US-08-659-251-47	Sequence 47, Appl	323	15	0.5	408	2	US-08-116-778B-19	Sequence 19, Appl
c 251	15	0.5	118	2	US-09-456-490-47	Sequence 47, Appl	324	15	0.5	408	2	US-08-454-680-5	Sequence 5, Appl
c 252	15	0.5	118	5	PCT-US96-11445-47	Sequence 47, Appl	325	15	0.5	408	2	US-08-438-562-19	Sequence 19, Appl
c 253	15	0.5	126	4	US-09-172-841-2	Sequence 2, Appl	326	15	0.5	408	2	US-08-483-528B-19	Sequence 19, Appl
c 254	15	0.5	180	6	5188961-2	Patent No. 5188961	327	15	0.5	408	3	US-08-673-799C-19	Sequence 19, Appl
c 255	15	0.5	193	4	US-09-189-060B-20	Sequence 20, Appl	328	15	0.5	411	3	US-08-554-840-24	Sequence 24, Appl
c 256	15	0.5	206	1	US-07-670-611-14	Sequence 14, Appl	329	15	0.5	411	3	US-08-554-840-25	Sequence 25, Appl
c 257	15	0.5	206	1	US-08-220-674-14	Sequence 14, Appl	330	15	0.5	411	3	US-08-554-840-27	Sequence 27, Appl
c 258	15	0.5	206	1	US-08-445-186-14	Sequence 14, Appl	331	15	0.5	424	3	US-08-589-939-8	Sequence 8, Appl
c 259	15	0.5	206	1	US-08-446-549-14	Sequence 14, Appl	332	15	0.5	434	4	US-08-961-810-15	Sequence 15, Appl
c 260	15	0.5	206	2	US-08-446-550-14	Sequence 14, Appl	333	15	0.5	434	4	US-08-352-902B-15	Sequence 15, Appl
c 261	15	0.5	235	4	US-09-172-108-8	Sequence 8, Appl	c 334	15	0.5	446	4	US-08-569-147-83	Sequence 83, Appl
c 262	15	0.5	249	4	US-09-438-938-2	Sequence 2, Appl	335	15	0.5	447	2	US-08-752-844-1	Sequence 1, Appl
c 263	15	0.5	265	3	US-09-109-205-9	Sequence 9, Appl	336	15	0.5	447	2	US-08-779-940-1	Sequence 1, Appl
c 264	15	0.5	276	4	US-09-206-903A-3	Sequence 3, Appl	337	15	0.5	447	2	US-08-591-196-1	Sequence 1, Appl
c 265	15	0.5	276	4	US-09-206-903A-10	Sequence 10, Appl	338	15	0.5	447	4	US-08-838-692-3	Sequence 3, Appl
c 266	15	0.5	276	4	US-09-202-122-3	Sequence 3, Appl	339	15	0.5	447	4	US-09-192-838B-1	Sequence 1, Appl
c 267	15	0.5	276	4	US-09-202-122-10	Sequence 10, Appl	340	15	0.5	458	4	US-08-961-810-16	Sequence 16, Appl
c 268	15	0.5	276	4	US-09-206-935-20	Sequence 20, Appl	341	15	0.5	458	4	US-08-352-902B-16	Sequence 16, Appl
c 269	15	0.5	276	4	US-09-206-935-21	Sequence 21, Appl	342	15	0.5	477	1	US-08-050-319B-53	Sequence 53, Appl
c 270	15	0.5	276	4	US-09-206-936-20	Sequence 20, Appl	343	15	0.5	477	2	US-08-465-982-53	Sequence 53, Appl
c 271	15	0.5	276	4	US-09-206-936-21	Sequence 21, Appl	344	15	0.5	483	2	US-08-561-521-1	Sequence 1, Appl
c 272	15	0.5	283	6	5188961-3	Patent No. 5188961	345	15	0.5	483	4	US-09-326-394-1	Sequence 1, Appl
c 273	15	0.5	301	4	US-09-439-313-257	Sequence 257, App	346	15	0.5	483	5	PCT-US95-01219-1	Sequence 1, Appl
c 274	15	0.5	307	2	US-08-691-814B-107	Sequence 107, App	c 347	15	0.5	491	4	US-09-103-875-15	Sequence 15, Appl
c 275	15	0.5	318	2	US-08-808-158-1	Sequence 1, Appl	348	15	0.5	501	1	US-08-050-319B-1	Sequence 1, Appl
c 276	15	0.5	321	1	US-08-211-202-4	Sequence 4, Appl	349	15	0.5	501	2	US-08-465-982-1	Sequence 1, Appl
c 277	15	0.5	321	3	US-08-783-853A-6	Sequence 6, Appl	350	15	0.5	504	1	US-08-050-319B-56	Sequence 56, Appl
c 278	15	0.5	321	3	US-08-483-749A-11	Sequence 11, Appl	351	15	0.5	504	2	US-08-465-982-56	Sequence 56, Appl
c 279	15	0.5	324	1	US-08-236-520-3	Sequence 3, Appl	c 352	15	0.5	509	4	US-09-385-982-43	Sequence 43, Appl
c 280	15	0.5	324	1	US-08-040-204-9	Sequence 9, Appl	353	15	0.5	551	1	US-08-076-091C-14	Sequence 14, Appl
c 281	15	0.5	324	1	US-08-040-204-11	Sequence 11, Appl	354	15	0.5	551	1	US-08-285-641-14	Sequence 14, Appl
c 282	15	0.5	324	1	US-08-040-204-12	Sequence 12, Appl	c 355	15	0.5	578	4	US-09-385-982-113	Sequence 113, App
c 283	15	0.5	324	1	US-08-040-204-13	Sequence 13, Appl	c 356	15	0.5	580	2	US-08-809-763-2	Sequence 2, Appl
c 284	15	0.5	324	1	US-08-040-204-14	Sequence 14, Appl	c 357	15	0.5	591	4	US-09-385-982-300	Sequence 300, App
c 285	15	0.5	324	1	US-08-040-204-15	Sequence 15, Appl	c 358	15	0.5	596	3	US-09-284-782-31	Sequence 31, Appl
c 286	15	0.5	324	2	US-08-040-204-16	Sequence 16, Appl	c 359	15	0.5	600	1	US-08-050-319B-47	Sequence 47, Appl
c 287	15	0.5	324	2	US-08-737-560A-13	Sequence 13, Appl	360	15	0.5	600	2	US-08-465-982-47	Sequence 47, Appl
c 288	15	0.5	324	3	US-09-065-059-4	Sequence 4, Appl	361	15	0.5	602	4	US-09-040-984-11	Sequence 11, Appl
c 289	15	0.5	324	3	US-09-247-352-14	Sequence 14, Appl	362	15	0.5	602	4	US-09-123-912-11	Sequence 11, Appl
c 290	15	0.5	324	5	PCT-US95-05262-3	Sequence 3, Appl	363	15	0.5	602	4	US-09-123-912-11	Sequence 11, Appl
c 291	15	0.5	327	2	US-08-553-497A-13	Sequence 13, Appl	364	15	0.5	606	4	US-09-040-984-55	Sequence 55, Appl
c 292	15	0.5	331	1	US-08-017-570-1	Sequence 1, Appl	365	15	0.5	606	4	US-09-123-912-55	Sequence 55, Appl
c 293	15	0.5	331	1	US-08-471-426-1	Sequence 1, Appl	c 366	15	0.5	622	1	US-07-620-426B-29	Sequence 29, Appl
c 294	15	0.5	331	5	PCT-US94-01709-1	Sequence 1, Appl	c 367	15	0.5	622	1	US-07-662-007B-35	Sequence 35, Appl
c 295	15	0.5	335	3	US-08-783-853A-102	Sequence 102, App	c 368	15	0.5	622	1	US-07-824-247-35	Sequence 35, Appl
c 296	15	0.5	336	3	US-09-065-059-14	Sequence 14, Appl	c 369	15	0.5	622	3	US-08-470-204A-35	Sequence 35, Appl
c 297	15	0.5	342	2	US-08-428-257A-75	Sequence 75, Appl	c 370	15	0.5	627	4	US-09-385-982-514	Sequence 514, App
c 298	15	0.5	342	3	US-08-929-856-65	Sequence 65, Appl	371	15	0.5	630	4	US-09-439-313-358	Sequence 358, App
c 299	15	0.5	342	3	US-08-929-856-188	Sequence 188, App	372	15	0.5	646	2	US-08-737-129A-3	Sequence 3, Appl
c 300	15	0.5	342	4	US-07-867-264-15	Sequence 15, Appl	373	15	0.5	649	4	US-09-385-982-336	Sequence 336, App
c 301	15	0.5	353	4	US-09-328-111-150	Sequence 150, App	374	15	0.5	656	4	US-08-943-731-129	Sequence 129, App
c 302	15	0.5	353	2	US-08-737-560A-8	Sequence 8, Appl	c 375	15	0.5	660	4	US-09-417-822-41	Sequence 41, App
c 303	15	0.5	372	1	US-08-050-319B-3	Sequence 3, Appl	376	15	0.5	670	4	US-09-098-789-6	Sequence 6, Appl
c 304	15	0.5	372	2	US-08-465-982-3	Sequence 3, Appl	c 377	15	0.5	690	4	US-09-091-219-8	Sequence 8, Appl
c 305	15	0.5	376	1	US-08-117-373-8	Sequence 8, Appl	c 378	15	0.5	693	1	US-08-322-805-3	Sequence 3, Appl
c 306	15	0.5	377	2	US-08-449-287-1	Sequence 1, Appl	379	15	0.5	693	2	US-08-429-499-3	Sequence 3, Appl
c 307	15	0.5	380	3	US-08-483-528B-90	Sequence 90, Appl	380	15	0.5	693	3	US-09-152-733-3	Sequence 3, Appl
c 308	15	0.5	381	3	US-08-673-799C-90	Sequence 90, Appl	381	15	0.5	699	1	US-08-225-989-3	Sequence 3, Appl
c 309	15	0.5	382	3	US-08-236-520-1	Sequence 1, Appl	382	15	0.5	699	1	US-08-570-923-3	Sequence 3, Appl
c 310	15	0.5	393	5	PCT-US95-05262-1	Sequence 1, Appl	383	15	0.5	699	1	US-08-580-014-3	Sequence 3, Appl
c 311	15	0.5	394	1	US-07-977-696C-10	Sequence 10, Appl	384	15	0.5	699	2	US-08-595-043A-49	Sequence 49, Appl
c 312	15	0.5	394	1	US-07-977-696C-64	Sequence 64, Appl	385	15	0.5	699	3	US-09-079-785-3	Sequence 3, Appl
c 313	15	0.5	394	1	US-08-129-930B-10	Sequence 10, Appl	386	15	0.5	699	4	US-09-522-217-33	Sequence 33, Appl
c 314	15	0.5	394	1	US-08-129-930B-64	Sequence 64, Appl	387	15	0.5	705	1	US-07-797-556-3	Sequence 3, Appl
c 315	15	0.5	394	4	US-08-976-288A-10	Sequence 10, Appl	388	15	0.5	705	1	US-08-308-881-3	Sequence 3, Appl
c 316	15	0.5	394	4	US-08-976-288A-64	Sequence 64, Appl	389	15	0.5	705	2	US-09-058-263-3	Sequence 3, Appl
c 317	15	0.5	396	2	US-08-483-636-1	Sequence 1, Appl	390	15	0.5	705	2	US-09-059-099-3	Sequence 3, Appl
c 318	15	0.5	396	2	US-08-483-632-1	Sequence 1, Appl	391	15	0.5	705	3	US-09-058-264-3	Sequence 3, Appl
c 319	15	0.5	406	2	US-08-561-521-14	Sequence 14, Appl	392	15	0.5	705	4	US-09-131-247-5	Sequence 5, Appl

393	15	0.5	705	5	PCT-US95-06530-3	Sequence 3, App1	466	15	0.5	1050	4	US-08-466-465-7	Sequence 7, App1
394	15	0.5	705	5	PCT-US95-15781-7	Sequence 7, App1	467	15	0.5	1050	5	PCT-US92-02050-42	Sequence 42, App1
395	15	0.5	708	1	US-08-488-376-18	Sequence 18, App1	468	15	0.5	1060	5	US-09-188-930-9	Sequence 9, App1
396	15	0.5	708	2	US-08-634-223-18	Sequence 18, App1	469	15	0.5	1080	4	US-09-180-100-14	Sequence 14, App1
397	15	0.5	708	2	US-08-634-224-18	Sequence 18, App1	470	15	0.5	1084	1	US-08-245-754-1	Sequence 1, App1
398	15	0.5	708	2	US-08-634-400-18	Sequence 18, App1	471	15	0.5	1094	2	US-08-597-731-1	Sequence 1, App1
399	15	0.5	708	2	US-08-635-878-18	Sequence 18, App1	472	15	0.5	1095	3	US-08-875-811-52	Sequence 52, App1
400	15	0.5	708	4	US-09-335-6978-18	Sequence 18, App1	473	15	0.5	1098	3	US-08-875-811-54	Sequence 54, App1
401	15	0.5	720	3	US-08-946-026-14	Sequence 14, App1	474	15	0.5	1125	2	US-08-583-5628-9	Sequence 9, App1
402	15	0.5	720	3	US-09-328-111-821	Sequence 821, App	475	15	0.5	1135	2	US-08-779-113-9	Sequence 9, App1
403	15	0.5	722	2	US-08-860-882A-26	Sequence 26, App1	476	15	0.5	1135	1	US-08-236-311-8	Sequence 8, App1
404	15	0.5	722	4	US-09-423-439-57	Sequence 57, App1	477	15	0.5	1135	4	US-08-457-918-8	Sequence 8, App1
405	15	0.5	723	4	US-09-227-357-1	Sequence 1, App1	478	15	0.5	1147	4	US-08-804-166-5	Sequence 5, App1
406	15	0.5	733	4	US-09-527-236A-27	Sequence 27, App1	479	15	0.5	1149	1	US-08-470-299-6	Sequence 6, App1
407	15	0.5	737	2	US-08-691-814B-116	Sequence 116, App	480	15	0.5	1149	1	US-08-470-299-9	Sequence 9, App1
408	15	0.5	740	1	US-08-360-823A-3	Sequence 3, App1	481	15	0.5	1150	4	US-08-811-481-33	Sequence 3, App1
409	15	0.5	740	2	US-08-249-189-3	Sequence 3, App1	482	15	0.5	1153	4	US-08-811-481-6	Sequence 6, App1
410	15	0.5	740	2	US-08-484-624A-3	Sequence 3, App1	483	15	0.5	1163	4	US-08-470-299-3	Sequence 3, App1
411	15	0.5	740	2	US-08-477-733B-3	Sequence 3, App1	484	15	0.5	1164	1	US-09-036-987A-32	Sequence 32, App1
412	15	0.5	740	3	US-08-763-995-3	Sequence 3, App1	485	15	0.5	1165	4	US-09-370-700-32	Sequence 32, App1
413	15	0.5	740	3	US-09-088-913A-3	Sequence 3, App1	486	15	0.5	1165	4	US-09-131-247-13	Sequence 13, App1
414	15	0.5	740	4	US-08-769-819-3	Sequence 3, App1	487	15	0.5	1170	4	US-09-131-247-15	Sequence 15, App1
415	15	0.5	740	4	US-08-770-974-3	Sequence 3, App1	488	15	0.5	1170	4	US-09-131-247-13	Sequence 13, App1
416	15	0.5	740	4	US-08-966-316-2	Sequence 2, App1	489	15	0.5	1172	2	US-08-070-901-21	Sequence 21, App1
417	15	0.5	745	1	US-07-943-843-7	Sequence 7, App1	490	15	0.5	1178	4	US-09-091-405-1	Sequence 1, App1
418	15	0.5	745	1	US-08-236-918A-14	Sequence 14, App1	491	15	0.5	1182	4	US-09-180-100-18	Sequence 18, App1
419	15	0.5	745	1	US-08-347-003-7	Sequence 7, App1	492	15	0.5	1197	2	US-08-530-165-6	Sequence 6, App1
420	15	0.5	745	4	US-09-150-864A-14	Sequence 14, App1	493	15	0.5	1202	4	US-08-804-166-3	Sequence 3, App1
421	15	0.5	751	2	US-08-822-260-2	Sequence 2, App1	494	15	0.5	1202	4	US-08-910-991-3	Sequence 3, App1
422	15	0.5	758	4	US-08-811-481-32	Sequence 32, App1	495	15	0.5	1216	3	US-08-808-720-4	Sequence 4, App1
423	15	0.5	758	4	US-09-218-950-32	Sequence 32, App1	496	15	0.5	1221	1	US-08-808-720-2	Sequence 2, App1
424	15	0.5	768	4	US-09-173-151A-11	Sequence 11, App1	497	15	0.5	1221	3	US-08-322-805-4	Sequence 4, App1
425	15	0.5	768	4	US-08-946-026-18	Sequence 18, App1	498	15	0.5	1231	2	US-08-429-499-4	Sequence 4, App1
426	15	0.5	800	3	US-08-477-877B-90	Sequence 90, App1	499	15	0.5	1231	3	US-09-152-733-3	Sequence 3, App1
427	15	0.5	807	1	US-08-472-281A-90	Sequence 90, App1	500	15	0.5	1234	3	US-08-808-720-6	Sequence 6, App1
428	15	0.5	807	1	US-08-477-877B-90	Sequence 90, App1	501	15	0.5	1234	4	US-09-397-992A-1	Sequence 1, App1
429	15	0.5	807	2	US-08-477-877B-90	Sequence 90, App1	502	15	0.5	1235	3	US-08-808-720-8	Sequence 8, App1
430	15	0.5	819	4	US-09-349-627-2	Sequence 2, App1	503	15	0.5	1269	5	PCT-US95-03866-13	Sequence 13, App1
431	15	0.5	836	2	US-08-933-616-1	Sequence 1, App1	504	15	0.5	1272	5	PCT-US95-03866-11	Sequence 11, App1
432	15	0.5	836	2	US-08-279-307-1	Sequence 1, App1	505	15	0.5	1280	5	PCT-US95-03866-13	Sequence 13, App1
433	15	0.5	847	2	US-08-773-423-4	Sequence 4, App1	506	15	0.5	1280	5	US-09-188-930-246	Sequence 246, App
434	15	0.5	847	2	US-08-332-419-3	Sequence 3, App1	507	15	0.5	1289	2	US-08-594-719C-3	Sequence 3, App1
435	15	0.5	882	1	US-08-207-481-19	Sequence 19, App1	508	15	0.5	1301	4	US-08-804-166-7	Sequence 7, App1
436	15	0.5	887	5	PCT-US95-02689-19	Sequence 19, App1	509	15	0.5	1301	4	US-08-804-166-7	Sequence 7, App1
437	15	0.5	894	4	US-08-811-481-3	Sequence 3, App1	510	15	0.5	1302	1	US-08-719-331-7	Sequence 7, App1
438	15	0.5	921	4	US-09-247-155-133	Sequence 133, App	511	15	0.5	1307	2	US-08-322-761A-1	Sequence 1, App1
439	15	0.5	921	4	US-08-401-908-26	Sequence 26, App1	512	15	0.5	1307	2	US-08-322-761A-1	Sequence 1, App1
440	15	0.5	931	4	US-09-247-155-140	Sequence 140, App	513	15	0.5	1307	4	US-09-307-912-1	Sequence 1, App1
441	15	0.5	935	2	US-08-832-028-73	Sequence 73, App1	514	15	0.5	1310	2	PCT-US95-04464-1	Sequence 1, App1
442	15	0.5	935	4	US-08-479-285-73	Sequence 73, App1	515	15	0.5	1317	1	US-08-394-152A-41	Sequence 41, App1
443	15	0.5	940	1	US-08-353-400-24	Sequence 24, App1	516	15	0.5	1317	1	US-08-097-827-10	Sequence 10, App1
444	15	0.5	942	4	US-08-713-556F-35	Sequence 35, App1	517	15	0.5	1317	1	US-08-370-975B-2	Sequence 2, App1
445	15	0.5	973	1	US-08-532-309-1	Sequence 1, App1	518	15	0.5	1317	1	US-08-464-517-10	Sequence 10, App1
446	15	0.5	974	1	US-09-178-869-3	Sequence 3, App1	519	15	0.5	1317	2	US-08-464-517-34	Sequence 34, App1
447	15	0.5	984	1	US-08-470-299-8	Sequence 8, App1	520	15	0.5	1317	2	US-08-246-361A-34	Sequence 34, App1
448	15	0.5	986	2	US-08-865-337A-4	Sequence 4, App1	521	15	0.5	1317	5	PCT-US93-05000-34	Sequence 34, App1
449	15	0.5	1005	1	US-08-436-463-5	Sequence 5, App1	522	15	0.5	1317	5	PCT-US93-05000-36	Sequence 36, App1
450	15	0.5	1005	1	US-08-470-299-5	Sequence 5, App1	523	15	0.5	1318	1	US-08-570-311-21	Sequence 21, App1
451	15	0.5	1006	1	US-09-198-955A-3	Sequence 3, App1	524	15	0.5	1320	6	5177197-2	Patent No. 5177197
452	15	0.5	1019	4	US-09-178-869-1	Sequence 1, App1	525	15	0.5	1329	5	PCT-US96-13155-3	Sequence 3, App1
453	15	0.5	1019	4	US-09-073-684-1	Sequence 7, App1	526	15	0.5	1343	3	US-08-718-738-3	Sequence 3, App1
454	15	0.5	1026	3	US-09-198-956-7	Sequence 7, App1	527	15	0.5	1343	4	US-09-221-844-3	Sequence 4, App1
455	15	0.5	1026	4	US-09-198-956-7	Sequence 7, App1	528	15	0.5	1347	2	US-08-822-830E-5	Sequence 5, App1
456	15	0.5	1026	4	US-09-395-858A-11	Sequence 11, App1	529	15	0.5	1347	2	US-08-850-660-5	Sequence 5, App1
457	15	0.5	1026	4	US-09-546-500-1	Sequence 1, App1	530	15	0.5	1350	1	US-08-157-101A-9	Sequence 9, App1
458	15	0.5	1026	4	US-09-546-762-1	Sequence 1, App1	531	15	0.5	1350	1	US-08-570-311-15	Sequence 15, App1
459	15	0.5	1049	4	US-08-804-166-1	Sequence 1, App1	532	15	0.5	1352	5	PCT-US95-03322A-3	Sequence 3, App1
460	15	0.5	1049	4	US-08-910-991-1	Sequence 1, App1	533	15	0.5	1353	4	US-09-028-274A-17	Sequence 17, App1
461	15	0.5	1050	1	US-07-940-861-42	Sequence 42, App1	534	15	0.5	1368	1	US-08-570-311-17	Sequence 17, App1
462	15	0.5	1050	1	US-08-459-517-42	Sequence 42, App1	535	15	0.5	1368	1	US-08-570-311-19	Sequence 19, App1
463	15	0.5	1050	1	US-08-459-517-42	Sequence 42, App1	536	15	0.5	1383	1	US-08-822-028-72	Sequence 72, App1
464	15	0.5	1050	2	US-08-460-132-42	Sequence 42, App1	537	15	0.5	1383	2	US-08-822-028-72	Sequence 72, App1
465	15	0.5	1050	2	US-08-460-132-42	Sequence 42, App1	538	15	0.5	1383	2	US-08-822-028-72	Sequence 72, App1

539	15	0.5	1383	4	US-08-479-285-72	Sequence 72, Appl	612	15	0.5	1644	2	US-08-792-824-11	Sequence 11, Appl
540	15	0.5	1383	4	US-08-479-285-72	Sequence 72, Appl	613	15	0.5	1645	2	US-08-461-812-1	Sequence 1, Appl
541	15	0.5	1392	4	US-09-257-584-1	Sequence 1, Appl	614	15	0.5	1655	2	US-09-049-672A-1	Sequence 21, Appl
542	15	0.5	1404	3	US-08-523-894-7	Sequence 7, Appl	615	15	0.5	1672	2	US-08-792-824-2	Sequence 2, Appl
543	15	0.5	1404	3	US-08-523-894-9	Sequence 9, Appl	616	15	0.5	1675	2	US-08-244-205-12	Sequence 12, Appl
544	15	0.5	1404	3	US-08-523-894-11	Sequence 11, Appl	617	15	0.5	1675	2	PCT-US92-10284-12	Sequence 12, Appl
545	15	0.5	1404	3	US-09-485-737B-66	Sequence 66, Appl	618	15	0.5	1683	1	US-07-945-283-3	Sequence 3, Appl
546	15	0.5	1416	4	US-09-330-611-9	Sequence 9, Appl	619	15	0.5	1683	1	US-09-192-545-1	Sequence 1, Appl
547	15	0.5	1416	4	US-08-793-450-7	Sequence 7, Appl	620	15	0.5	1716	2	US-08-822-028-15	Sequence 15, Appl
548	15	0.5	1425	2	US-08-249-189-15	Sequence 15, Appl	621	15	0.5	1716	2	US-08-479-285-15	Sequence 15, Appl
549	15	0.5	1425	2	US-08-484-624A-15	Sequence 15, Appl	622	15	0.5	1737	4	US-09-173-151A-1	Sequence 1, Appl
550	15	0.5	1425	2	US-08-477-733B-15	Sequence 15, Appl	623	15	0.5	1743	3	US-08-665-259-20	Sequence 20, Appl
551	15	0.5	1425	3	US-09-088-913A-15	Sequence 15, Appl	624	15	0.5	1743	3	US-08-762-500-20	Sequence 20, Appl
552	15	0.5	1425	4	US-08-769-819-15	Sequence 15, Appl	625	15	0.5	1747	1	US-08-203-389-1	Sequence 1, Appl
553	15	0.5	1425	4	US-08-770-974-15	Sequence 15, Appl	626	15	0.5	1748	1	US-08-202-056-8	Sequence 8, Appl
554	15	0.5	1428	1	US-08-488-376-17	Sequence 17, Appl	627	15	0.5	1748	3	US-09-100-730-1	Sequence 1, Appl
555	15	0.5	1428	1	US-08-488-376-19	Sequence 19, Appl	628	15	0.5	1794	4	US-08-123-938A-9	Sequence 9, Appl
556	15	0.5	1428	2	US-08-634-223-17	Sequence 17, Appl	629	15	0.5	1794	4	PCT-US94-10080-9	Sequence 9, Appl
557	15	0.5	1428	2	US-08-634-223-19	Sequence 19, Appl	630	15	0.5	1796	3	US-08-477-460B-1	Sequence 1, Appl
558	15	0.5	1428	2	US-08-634-224-17	Sequence 17, Appl	631	15	0.5	1796	3	US-08-379-516-1	Sequence 1, Appl
559	15	0.5	1428	2	US-08-634-224-19	Sequence 19, Appl	632	15	0.5	1796	4	US-09-329-916-1	Sequence 1, Appl
560	15	0.5	1428	2	US-08-634-400-17	Sequence 17, Appl	633	15	0.5	1796	4	US-08-485-372A-1	Sequence 1, Appl
561	15	0.5	1428	2	US-08-634-400-19	Sequence 19, Appl	634	15	0.5	1796	4	US-09-409-006A-1	Sequence 1, Appl
562	15	0.5	1428	2	US-08-635-878-17	Sequence 17, Appl	635	15	0.5	1796	5	PCT-US93-07422-1	Sequence 5, Appl
563	15	0.5	1428	2	US-08-635-878-19	Sequence 19, Appl	636	15	0.5	1804	1	US-08-306-691B-40	Sequence 40, Appl
564	15	0.5	1428	2	US-08-770-057-17	Sequence 17, Appl	637	15	0.5	1804	5	PCT-US93-06251-82	Sequence 82, Appl
565	15	0.5	1428	2	US-08-770-057-19	Sequence 19, Appl	638	15	0.5	1833	4	US-09-173-151A-13	Sequence 13, Appl
566	15	0.5	1428	4	US-09-335-697B-17	Sequence 17, Appl	639	15	0.5	1845	2	US-08-622-740-5	Sequence 5, Appl
567	15	0.5	1428	4	US-09-335-697B-19	Sequence 19, Appl	640	15	0.5	1845	3	US-08-440-688-5	Sequence 5, Appl
568	15	0.5	1428	4	US-08-812-946A-3	Sequence 3, Appl	641	15	0.5	1845	4	US-09-122-399-5	Sequence 5, Appl
569	15	0.5	1431	3	US-08-487-550-11	Sequence 3, Appl	642	15	0.5	1848	2	US-08-622-740-7	Sequence 7, Appl
570	15	0.5	1431	3	US-08-487-550-13	Sequence 11, Appl	643	15	0.5	1848	3	US-08-440-689-7	Sequence 7, Appl
571	15	0.5	1437	3	US-08-487-550-17	Sequence 7, Appl	644	15	0.5	1848	4	US-09-122-399-7	Sequence 7, Appl
572	15	0.5	1459	1	US-08-094-534-8	Sequence 8, Appl	645	15	0.5	1870	1	US-08-592-128-86	Sequence 86, Appl
573	15	0.5	1459	1	US-08-581-543-8	Sequence 8, Appl	646	15	0.5	1887	4	US-09-201-641-5	Sequence 5, Appl
574	15	0.5	1459	2	US-08-581-543-8	Sequence 8, Appl	647	15	0.5	1894	5	PCT-US96-10043-10	Sequence 10, Appl
575	15	0.5	1459	2	US-08-581-543-8	Sequence 8, Appl	648	15	0.5	1894	5	PCT-US96-10043-10	Sequence 10, Appl
576	15	0.5	1459	5	PCT-US94-08000-8	Sequence 8, Appl	649	15	0.5	1941	2	US-09-008-960-2	Sequence 2, Appl
577	15	0.5	1478	4	US-09-149-922-6	Sequence 6, Appl	650	15	0.5	1941	3	US-09-368-240-2	Sequence 2, Appl
578	15	0.5	1480	4	US-09-149-922-6	Sequence 6, Appl	651	15	0.5	1947	4	US-09-468-702-2	Sequence 2, Appl
579	15	0.5	1480	4	US-09-290-640-65	Sequence 65, Appl	652	15	0.5	1947	1	US-08-299-848B-19	Sequence 19, Appl
580	15	0.5	1494	2	US-08-583-562B-11	Sequence 11, Appl	653	15	0.5	1947	2	US-08-142-368A-19	Sequence 19, Appl
581	15	0.5	1494	2	US-08-779-113-11	Sequence 11, Appl	654	15	0.5	1947	3	US-08-967-727-19	Sequence 19, Appl
582	15	0.5	1502	4	US-09-206-903A-2	Sequence 2, Appl	655	15	0.5	1947	4	US-08-037-230D-19	Sequence 19, Appl
583	15	0.5	1502	4	US-09-206-903A-11	Sequence 11, Appl	656	15	0.5	1966	1	US-08-197-792-44	Sequence 44, Appl
584	15	0.5	1502	4	US-09-202-122-2	Sequence 2, Appl	657	15	0.5	1966	1	US-08-459-850-44	Sequence 44, Appl
585	15	0.5	1502	4	US-09-202-122-11	Sequence 11, Appl	658	15	0.5	1966	1	US-08-459-214-44	Sequence 44, Appl
586	15	0.5	1502	4	US-09-206-935-2	Sequence 2, Appl	659	15	0.5	1974	3	US-08-762-500-78	Sequence 78, Appl
587	15	0.5	1502	4	US-09-206-935-3	Sequence 3, Appl	660	15	0.5	1980	1	US-08-109-106-4	Sequence 4, Appl
588	15	0.5	1502	4	US-09-206-936-2	Sequence 2, Appl	661	15	0.5	1983	3	US-08-109-106-4	Sequence 4, Appl
589	15	0.5	1510	1	US-07-759-568-4	Sequence 4, Appl	662	15	0.5	1983	3	US-09-181-706-7	Sequence 7, Appl
590	15	0.5	1510	1	US-08-466-337A-17	Sequence 17, Appl	663	15	0.5	1983	4	US-09-458-791-7	Sequence 7, Appl
591	15	0.5	1550	2	US-08-475-359-17	Sequence 17, Appl	664	15	0.5	1983	4	US-09-459-066-7	Sequence 7, Appl
592	15	0.5	1550	2	US-08-802-322-2	Sequence 2, Appl	665	15	0.5	1995	3	US-08-957-063-15	Sequence 15, Appl
593	15	0.5	1550	3	US-08-802-322-2	Sequence 2, Appl	666	15	0.5	1995	3	US-08-957-063-15	Sequence 15, Appl
594	15	0.5	1550	3	US-08-895-601-3	Sequence 3, Appl	667	15	0.5	1995	4	US-09-487-683-17	Sequence 17, Appl
595	15	0.5	1557	1	US-08-385-229-3	Sequence 3, Appl	668	15	0.5	2004	2	US-08-691-814B-1	Sequence 1, Appl
596	15	0.5	1567	3	US-09-049-672A-17	Sequence 17, Appl	669	15	0.5	2009	1	US-08-109-106-5	Sequence 5, Appl
597	15	0.5	1567	3	US-08-157-101A-6	Sequence 6, Appl	670	15	0.5	2009	1	US-08-109-106-5	Sequence 5, Appl
598	15	0.5	1576	1	US-08-477-254A-6	Sequence 6, Appl	671	15	0.5	2010	3	US-09-070-637-19	Sequence 19, Appl
599	15	0.5	1587	2	US-08-428-576B-6	Sequence 6, Appl	672	15	0.5	2029	2	US-07-916-098A-43	Sequence 43, Appl
600	15	0.5	1587	2	US-08-428-576B-6	Sequence 6, Appl	673	15	0.5	2038	1	US-08-181-271A-1	Sequence 1, Appl
601	15	0.5	1587	2	US-08-428-576B-6	Sequence 6, Appl	674	15	0.5	2038	1	US-08-449-803-1	Sequence 1, Appl
602	15	0.5	1587	2	US-08-428-576B-6	Sequence 6, Appl	675	15	0.5	2038	1	US-08-449-803-1	Sequence 1, Appl
603	15	0.5	1587	2	US-08-428-576B-6	Sequence 6, Appl	676	15	0.5	2038	1	US-08-449-803-1	Sequence 1, Appl
604	15	0.5	1591	2	US-08-646-981-14	Sequence 14, Appl	677	15	0.5	2038	1	US-08-449-803-1	Sequence 1, Appl
605	15	0.5	1594	4	US-07-934-393B-3	Sequence 3, Appl	678	15	0.5	2038	1	US-08-456-265A-1	Sequence 1, Appl
606	15	0.5	1601	2	US-08-838-957A-3	Sequence 3, Appl	679	15	0.5	2038	1	US-08-455-244-1	Sequence 1, Appl
607	15	0.5	1617	2	US-08-378-939-9	Sequence 9, Appl	680	15	0.5	2038	1	US-08-454-876-1	Sequence 1, Appl
608	15	0.5	1617	2	US-08-378-939-9	Sequence 9, Appl	681	15	0.5	2038	1	US-08-457-364-1	Sequence 1, Appl
609	15	0.5	1632	2	US-08-792-824-8	Sequence 8, Appl	682	15	0.5	2038	2	US-08-457-364-1	Sequence 1, Appl
610	15	0.5	1641	2	US-08-792-824-8	Sequence 8, Appl	683	15	0.5	2038	2	US-08-456-262-1	Sequence 1, Appl
611	15	0.5	1642	3	US-08-923-230-1	Sequence 1, Appl	684	15	0.5	2038	2	US-08-456-262-1	Sequence 1, Appl

685	15	0.5	2038	2	US-08-455-736-1	Sequence 1, Appl1	c 758	15	0.5	2576	3	US-09-300-529-35	Sequence 35, Appl1
686	15	0.5	2038	2	US-08-971-217-1	Sequence 1, Appl1	c 759	15	0.5	2580	3	US-09-050-863-2	Sequence 2, Appl1
587	15	0.5	2038	4	US-09-350-600-1	Sequence 1, Appl1	c 760	15	0.5	2580	4	US-09-359-081-2	Sequence 2, Appl1
588	15	0.5	2043	1	US-07-737-736B-6	Sequence 6, Appl1	c 761	15	0.5	2582	2	US-08-816-105A-2	Sequence 2, Appl1
589	15	0.5	2043	3	US-08-227-496C-14	Sequence 1, Appl1	c 762	15	0.5	2582	2	US-08-816-105A-2	Sequence 2, Appl1
590	15	0.5	2061	4	US-09-173-151A-3	Sequence 3, Appl1	c 763	15	0.5	2652	1	US-08-106-433A-17	Sequence 1, Appl1
691	15	0.5	2062	1	US-08-050-319B-24	Sequence 24, Appl1	c 764	15	0.5	2655	1	US-08-471-033-17	Sequence 17, Appl1
692	15	0.5	2062	2	US-08-465-982-24	Sequence 2, Appl1	c 765	15	0.5	2655	1	US-08-471-033-17	Sequence 26, Appl1
693	15	0.5	2067	4	US-09-058-368-5	Sequence 5, Appl1	c 766	15	0.5	2655	2	US-08-471-044-17	Sequence 17, Appl1
694	15	0.5	2133	3	US-08-947-965-1	Sequence 1, Appl1	c 767	15	0.5	2655	2	US-08-471-044-17	Sequence 26, Appl1
695	15	0.5	2133	4	US-09-485-737B-89	Sequence 8, Appl1	c 768	15	0.5	2655	2	US-08-463-483A-17	Sequence 17, Appl1
c 696	15	0.5	2136	3	US-08-946-475-8	Sequence 8, Appl1	c 769	15	0.5	2655	2	US-08-471-046A-17	Sequence 26, Appl1
c 697	15	0.5	2136	4	US-09-340-479-8	Sequence 8, Appl1	c 770	15	0.5	2655	2	US-08-471-046A-17	Sequence 17, Appl1
698	15	0.5	2150	1	US-08-580-680-2	Sequence 2, Appl1	c 771	15	0.5	2655	2	US-08-470-556B-17	Sequence 26, Appl1
699	15	0.5	2150	1	US-08-480-156A-2	Sequence 2, Appl1	c 772	15	0.5	2655	2	US-08-470-556B-17	Sequence 17, Appl1
700	15	0.5	2150	1	US-08-354-961-2	Sequence 2, Appl1	c 773	15	0.5	2655	2	US-08-469-334-17	Sequence 26, Appl1
701	15	0.5	2150	2	US-08-581-094-2	Sequence 2, Appl1	c 774	15	0.5	2655	2	US-08-469-334-17	Sequence 17, Appl1
702	15	0.5	2150	2	US-08-580-665-2	Sequence 2, Appl1	c 775	15	0.5	2655	3	US-09-300-529-17	Sequence 26, Appl1
703	15	0.5	2150	2	US-08-581-142-2	Sequence 2, Appl1	c 776	15	0.5	2655	3	US-09-300-529-17	Sequence 26, Appl1
704	15	0.5	2150	3	US-08-581-103-2	Sequence 2, Appl1	c 777	15	0.5	2659	1	US-08-594-031-91	Sequence 91, Appl1
705	15	0.5	2150	5	PCT-US94-05442A-2	Sequence 2, Appl1	c 778	15	0.5	2659	1	US-08-594-031-91	Sequence 101, Appl1
706	15	0.5	2151	3	US-08-964-449-3	Sequence 3, Appl1	c 779	15	0.5	2658	3	US-08-946-475-1	Sequence 1, Appl1
707	15	0.5	2161	3	US-09-106-038A-1	Sequence 1, Appl1	c 780	15	0.5	2658	4	US-09-340-479-1	Sequence 1, Appl1
708	15	0.5	2161	4	US-09-505-250-3	Sequence 3, Appl1	c 781	15	0.5	2722	3	US-08-804-439A-10	Sequence 10, Appl1
c 709	15	0.5	2166	4	US-08-317-401E-3	Sequence 3, Appl1	c 782	15	0.5	2722	3	US-08-720-229-10	Sequence 10, Appl1
710	15	0.5	2171	4	US-08-811-481-1	Sequence 1, Appl1	c 783	15	0.5	2726	4	US-08-811-481-23	Sequence 23, Appl1
711	15	0.5	2175	1	US-08-321-668-1	Sequence 1, Appl1	c 784	15	0.5	2808	4	US-08-237-919-1	Sequence 1, Appl1
712	15	0.5	2175	1	US-08-837-941-1	Sequence 1, Appl1	c 785	15	0.5	2808	4	US-08-732-429-1	Sequence 1, Appl1
713	15	0.5	2192	3	US-08-126-016-1	Sequence 1, Appl1	c 786	15	0.5	2808	4	US-08-732-429-1	Sequence 1, Appl1
714	15	0.5	2192	3	US-08-942-001-1	Sequence 1, Appl1	c 787	15	0.5	2840	5	PCT-US93-05521-64	Sequence 64, Appl1
715	15	0.5	2192	4	US-09-337-386-1	Sequence 1, Appl1	c 788	15	0.5	2840	5	PCT-US93-05521-64	Sequence 64, Appl1
716	15	0.5	2196	2	US-08-865-337B-2	Sequence 2, Appl1	c 789	15	0.5	2866	2	US-08-687-080-55	Sequence 55, Appl1
717	15	0.5	2235	1	US-08-313-181-3	Sequence 3, Appl1	c 790	15	0.5	2917	2	US-08-437-607B-3	Sequence 3, Appl1
c 718	15	0.5	2236	2	US-08-484-933B-13	Sequence 13, Appl1	c 791	15	0.5	2917	2	US-08-548-159-2	Sequence 2, Appl1
c 719	15	0.5	2236	2	US-08-484-158B-13	Sequence 13, Appl1	c 792	15	0.5	2974	1	US-08-290-978A-4	Sequence 4, Appl1
c 720	15	0.5	2236	2	US-08-484-596A-13	Sequence 13, Appl1	c 793	15	0.5	2974	2	US-08-780-869-4	Sequence 4, Appl1
c 721	15	0.5	2236	2	US-08-480-150A-13	Sequence 13, Appl1	c 794	15	0.5	2991	1	US-08-872-094-1	Sequence 1, Appl1
c 722	15	0.5	2236	3	US-08-458-731-13	Sequence 13, Appl1	c 795	15	0.5	2994	2	US-08-548-159-4	Sequence 4, Appl1
c 723	15	0.5	2236	3	US-08-449-223A-13	Sequence 13, Appl1	c 796	15	0.5	3046	1	US-08-276-099A-1	Sequence 1, Appl1
c 724	15	0.5	2259	4	US-09-173-151A-15	Sequence 15, Appl1	c 797	15	0.5	3046	4	US-08-781-890-1	Sequence 1, Appl1
c 725	15	0.5	2287	1	US-08-480-036-1	Sequence 1, Appl1	c 798	15	0.5	3046	4	US-08-781-890-1	Sequence 1, Appl1
c 726	15	0.5	2287	1	US-08-461-968A-1	Sequence 1, Appl1	c 799	15	0.5	3061	4	US-09-147-119-6	Sequence 11, Appl1
c 727	15	0.5	2287	2	US-08-462-571-1	Sequence 1, Appl1	c 800	15	0.5	3268	3	US-09-356-952-13	Sequence 6, Appl1
c 728	15	0.5	2287	5	PCT-US96-10043-8	Sequence 8, Appl1	c 801	15	0.5	3282	1	US-08-276-852-154	Sequence 154, Appl1
c 729	15	0.5	2287	5	PCT-US96-10043-8	Sequence 8, Appl1	c 802	15	0.5	3282	1	US-08-276-852-154	Sequence 154, Appl1
c 730	15	0.5	2294	3	US-08-964-700A-1	Sequence 1, Appl1	c 803	15	0.5	3282	1	US-08-276-852-154	Sequence 159, Appl1
c 731	15	0.5	2394	4	US-09-414-010-3	Sequence 3, Appl1	c 804	15	0.5	3282	1	US-08-276-852-154	Sequence 159, Appl1
c 732	15	0.5	2399	2	US-08-070-116A-1	Sequence 1, Appl1	c 805	15	0.5	3282	1	US-08-276-852-154	Sequence 159, Appl1
c 733	15	0.5	2411	3	US-09-188-930-75	Sequence 75, Appl1	c 806	15	0.5	3282	1	US-08-899-575-154	Sequence 154, Appl1
c 734	15	0.5	2411	3	US-09-188-930-75	Sequence 256, App	c 807	15	0.5	3282	1	US-08-899-575-154	Sequence 154, Appl1
c 735	15	0.5	2455	1	US-08-073-807A-1	Sequence 1, Appl1	c 808	15	0.5	3282	1	US-08-899-575-154	Sequence 159, Appl1
c 736	15	0.5	2464	4	US-08-811-481-21	Sequence 21, Appl1	c 809	15	0.5	3282	1	US-08-899-575-154	Sequence 154, Appl1
c 737	15	0.5	2482	3	US-08-477-460B-3	Sequence 3, Appl1	c 810	15	0.5	3282	1	US-08-899-575-154	Sequence 154, Appl1
c 738	15	0.5	2482	3	US-08-379-516-3	Sequence 3, Appl1	c 811	15	0.5	3282	1	US-08-899-575-154	Sequence 154, Appl1
c 739	15	0.5	2482	4	US-09-329-916-3	Sequence 3, Appl1	c 812	15	0.5	3282	1	US-08-899-575-154	Sequence 159, Appl1
c 740	15	0.5	2482	4	US-08-485-372A-3	Sequence 3, Appl1	c 813	15	0.5	3282	1	US-08-899-575-154	Sequence 159, Appl1
c 741	15	0.5	2482	4	US-09-409-006A-3	Sequence 3, Appl1	c 814	15	0.5	3282	5	PCT-US95-08743-154	Sequence 154, App
c 742	15	0.5	2484	5	PCT-US93-07422-3	Sequence 3, Appl1	c 815	15	0.5	3282	5	PCT-US95-08743-154	Sequence 154, App
c 743	15	0.5	2484	2	US-08-629-291A-32	Sequence 32, Appl1	c 816	15	0.5	3282	5	PCT-US95-08743-154	Sequence 159, App
c 744	15	0.5	2484	2	US-08-629-291A-32	Sequence 32, Appl1	c 817	15	0.5	3287	3	US-09-136-652-1	Sequence 169, App
c 745	15	0.5	2499	1	US-08-658-335B-32	Sequence 32, Appl1	c 818	15	0.5	3287	3	US-09-136-652-1	Sequence 169, App
c 746	15	0.5	2499	1	US-08-431-560-2	Sequence 2, Appl1	c 819	15	0.5	3287	1	US-08-920-812-8	Sequence 8, Appl1
c 747	15	0.5	2560	2	US-08-463-345-2	Sequence 2, Appl1	c 820	15	0.5	3287	1	US-08-920-812-8	Sequence 8, Appl1
c 748	15	0.5	2574	2	US-07-916-098A-44	Sequence 44, Appl1	c 821	15	0.5	3287	1	US-08-921-177-8	Sequence 8, Appl1
c 749	15	0.5	2574	2	US-08-583-562B-1	Sequence 1, Appl1	c 822	15	0.5	3287	2	US-08-362-577C-8	Sequence 8, Appl1
c 750	15	0.5	2574	2	US-08-583-562B-1	Sequence 1, Appl1	c 823	15	0.5	3287	2	US-08-362-577C-8	Sequence 8, Appl1
c 751	15	0.5	2574	2	US-08-779-113-1	Sequence 1, Appl1	c 824	15	0.5	3287	4	US-08-820-826-8	Sequence 8, Appl1
c 752	15	0.5	2576	2	US-08-779-113-1	Sequence 1, Appl1	c 825	15	0.5	3293	4	US-08-820-826-8	Sequence 15, Appl1
c 753	15	0.5	2576	2	US-08-471-033-35	Sequence 35, Appl1	c 826	15	0.5	3321	4	US-09-030-096D-1	Sequence 1, Appl1
c 754	15	0.5	2576	2	US-08-471-044-35	Sequence 35, Appl1	c 827	15	0.5	3310	4	US-08-484-438-5	Sequence 5, Appl1
c 755	15	0.5	2576	2	US-08-463-483A-35	Sequence 35, Appl1	c 828	15	0.5	3410	4	US-09-020-935-110	Sequence 110, App
c 756	15	0.5	2576	2	US-08-471-046A-35	Sequence 35, Appl1	c 829	15	0.5	3410	4	US-09-030-607-110	Sequence 110, App
c 757	15	0.5	2576	2	US-08-470-566B-35	Sequence 35, Appl1	c 830	15	0.5	3416	2	US-09-439-313-110	Sequence 110, App
c 758	15	0.5	2576	2	US-08-469-334-35	Sequence 35, Appl1	c 830	15	0.5	3416	2	US-08-357-642A-2	Sequence 2, Appl1

C 831	15	0.5	3416	2	US-08-460-626-2	Sequence 2, App11	C 904	15	0.5	4104	1	US-08-268-161A-94	Sequence 94, App1
C 832	15	0.5	3475	4	US-09-657-481A-10	Sequence 47, App1	C 905	15	0.5	4104	2	US-08-453-702A-94	Sequence 94, App1
C 833	15	0.5	3476	3	US-08-630-916A-47	Sequence 47, App1	C 906	15	0.5	4104	2	US-09-099-639-94	Sequence 94, App1
C 834	15	0.5	3480	1	US-07-657-769B-68	Sequence 68, App1	C 907	15	0.5	4104	5	PCT-US93-1256B-94	Sequence 94, App1
C 835	15	0.5	3480	1	US-07-789-184-219	Sequence 219, App	C 908	15	0.5	4104	5	PCT-US85-08071-94	Sequence 94, App1
C 836	15	0.5	3480	1	US-08-475-263-219	Sequence 219, App	C 909	15	0.5	4138	1	US-08-323-474-1	Sequence 1, App11
C 837	15	0.5	3480	1	US-08-485-886-219	Sequence 219, App	C 910	15	0.5	4138	1	US-08-447-411-75	Sequence 75, App1
C 838	15	0.5	3480	2	US-08-477-362-219	Sequence 219, App	C 911	15	0.5	4138	2	US-08-662-227-33	Sequence 33, App1
C 839	15	0.5	3480	2	US-08-477-134-219	Sequence 219, App	C 912	15	0.5	4138	4	US-09-017-947-33	Sequence 33, App1
C 840	15	0.5	3480	3	US-08-473-489A-219	Sequence 219, App	C 913	15	0.5	4138	5	PCT-US93-06093-1	Sequence 1, App11
C 841	15	0.5	3480	3	US-08-485-695-219	Sequence 219, App	C 914	15	0.5	4175	1	US-07-934-393B-1	Sequence 1, App11
C 842	15	0.5	3480	4	US-08-018-760-219	Sequence 219, App	C 915	15	0.5	4175	1	US-08-278-088A-1	Sequence 1, App11
C 843	15	0.5	3494	4	US-09-139-802-200	Sequence 200, App	C 916	15	0.5	4175	2	US-08-838-957A-1	Sequence 1, App11
C 844	15	0.5	3494	4	US-08-976-259-84	Sequence 84, App1	C 917	15	0.5	4176	2	US-08-878-089A-5	Sequence 5, App11
C 845	15	0.5	3520	1	US-08-424-788-1	Sequence 1, App11	C 918	15	0.5	4176	2	US-08-838-957A-5	Sequence 5, App11
C 846	15	0.5	3520	1	US-08-110-683-3	Sequence 3, App11	C 919	15	0.5	4248	1	US-08-678-614-1	Sequence 1, App11
C 847	15	0.5	3520	2	US-08-477-166-3	Sequence 3, App11	C 920	15	0.5	4252	2	US-08-475-844-1	Sequence 4, App11
C 848	15	0.5	3520	2	US-08-472-097-3	Sequence 3, App11	C 921	15	0.5	4252	5	PCT-US85-08429-4	Sequence 4, App11
C 849	15	0.5	3520	2	PCT-US93-11638-3	Sequence 3, App11	C 922	15	0.5	4268	1	US-08-264-534-33	Sequence 33, App1
C 850	15	0.5	3597	4	US-09-095-758-4	Sequence 4, App11	C 923	15	0.5	4268	1	US-08-083-590A-12	Sequence 12, App1
C 851	15	0.5	3597	4	US-09-422-968-4	Sequence 4, App11	C 924	15	0.5	4268	1	US-08-465-500-33	Sequence 33, App1
C 852	15	0.5	3642	3	US-08-946-026-16	Sequence 16, App1	C 925	15	0.5	4268	2	US-08-346-128-33	Sequence 33, App1
C 853	15	0.5	3747	2	US-09-080-897-1	Sequence 1, App11	C 926	15	0.5	4268	3	US-08-532-384-12	Sequence 12, App1
C 854	15	0.5	3747	4	US-09-323-735-1	Sequence 1, App11	C 927	15	0.5	4268	3	US-08-893-828-33	Sequence 33, App1
C 855	15	0.5	3767	4	US-08-928-941D-28	Sequence 28, App1	C 928	15	0.5	4328	1	US-08-322-742-12	Sequence 12, App1
C 856	15	0.5	3767	4	US-08-928-941D-30	Sequence 30, App1	C 929	15	0.5	4338	2	US-08-446-363-1	Sequence 1, App11
C 857	15	0.5	3767	4	US-09-280-590A-28	Sequence 28, App1	C 930	15	0.5	4338	2	US-08-015-986A-1	Sequence 1, App11
C 858	15	0.5	3767	4	US-09-280-590A-30	Sequence 30, App1	C 931	15	0.5	4378	2	US-09-080-897-3	Sequence 3, App11
C 859	15	0.5	3807	2	US-08-816-755-1	Sequence 1, App11	C 932	15	0.5	4378	2	US-09-323-735-3	Sequence 3, App11
C 860	15	0.5	3807	4	US-09-090-673-1	Sequence 1, App11	C 933	15	0.5	4379	1	US-08-592-214A-17	Sequence 17, App1
C 861	15	0.5	3865	1	US-08-832-883-48	Sequence 48, App1	C 934	15	0.5	4379	3	US-09-149-976-17	Sequence 17, App1
C 862	15	0.5	3865	2	US-08-832-877-48	Sequence 48, App1	C 935	15	0.5	4399	3	US-08-899-595-12	Sequence 2, App11
C 863	15	0.5	3867	3	US-08-762-428A-5	Sequence 5, App11	C 936	15	0.5	4435	2	US-08-792-824-1	Sequence 1, App11
C 864	15	0.5	3885	3	US-08-872-094-9	Sequence 9, App11	C 937	15	0.5	4486	1	US-08-322-742-15	Sequence 15, App1
C 865	15	0.5	3923	2	US-08-199-485-2	Sequence 2, App11	C 938	15	0.5	4510	2	US-08-570-311-1	Sequence 1, App11
C 866	15	0.5	3924	1	US-08-395-246C-1	Sequence 1, App11	C 939	15	0.5	4510	2	US-08-353-485-1	Sequence 1, App11
C 867	15	0.5	3934	1	US-09-226-568-18	Sequence 18, App1	C 940	15	0.5	4517	5	PCT-US93-06251-83	Sequence 83, App1
C 868	15	0.5	3946	1	US-08-077-848A-1	Sequence 1, App11	C 941	15	0.5	4601	3	US-08-470-058-3	Sequence 3, App11
C 869	15	0.5	3946	3	US-09-211-640-1	Sequence 1, App11	C 942	15	0.5	4601	3	US-09-037-188-3	Sequence 3, App11
C 870	15	0.5	3946	4	US-09-378-536-1	Sequence 1, App11	C 943	15	0.5	4601	3	US-09-285-310-3	Sequence 3, App11
C 871	15	0.5	3946	5	PCT-US94-03547-1	Sequence 1, App11	C 944	15	0.5	4610	4	US-09-417-822-4	Sequence 4, App11
C 872	15	0.5	3981	2	US-08-955-138-2	Sequence 2, App11	C 945	15	0.5	4622	4	US-08-509-024-6	Sequence 6, App11
C 873	15	0.5	4016	3	US-08-762-428A-7	Sequence 7, App11	C 946	15	0.5	4622	4	US-09-333-279-6	Sequence 6, App11
C 874	15	0.5	4030	4	US-09-293-505-1	Sequence 1, App11	C 947	15	0.5	4650	1	US-07-998-003A-102	Sequence 102, App
C 875	15	0.5	4031	1	US-08-471-033-49	Sequence 49, App1	C 948	15	0.5	4650	1	US-08-453-274A-102	Sequence 102, App
C 876	15	0.5	4031	2	US-08-471-044-49	Sequence 49, App1	C 949	15	0.5	4650	1	US-08-453-695A-102	Sequence 102, App
C 877	15	0.5	4031	2	US-08-463-483A-49	Sequence 49, App1	C 950	15	0.5	4650	1	US-08-268-161A-102	Sequence 102, App
C 878	15	0.5	4031	2	US-08-471-046A-49	Sequence 49, App1	C 951	15	0.5	4650	2	US-08-453-702A-102	Sequence 102, App
C 879	15	0.5	4031	2	US-08-470-566B-49	Sequence 49, App1	C 952	15	0.5	4650	4	US-09-099-639-102	Sequence 102, App
C 880	15	0.5	4031	3	US-08-469-334-49	Sequence 49, App1	C 953	15	0.5	4650	5	PCT-US93-1258B-102	Sequence 102, App
C 881	15	0.5	4031	3	US-09-300-529-49	Sequence 49, App1	C 954	15	0.5	4650	5	PCT-US93-08071-102	Sequence 102, App
C 882	15	0.5	4032	1	US-08-126-587C-8	Sequence 8, App11	C 955	15	0.5	4651	4	US-09-417-822-5	Sequence 3, App11
C 883	15	0.5	4060	1	US-08-164-292B-1	Sequence 1, App11	C 956	15	0.5	4651	4	US-09-417-822-5	Sequence 3, App11
C 884	15	0.5	4060	1	US-08-164-292B-3	Sequence 3, App11	C 957	15	0.5	4652	2	US-07-861-800-1	Sequence 1, App11
C 885	15	0.5	4060	1	US-08-164-292B-5	Sequence 5, App11	C 958	15	0.5	4652	2	US-07-861-800-1	Sequence 1, App11
C 886	15	0.5	4060	1	US-08-164-292B-7	Sequence 7, App11	C 959	15	0.5	4652	2	US-08-462-437-27	Sequence 27, App1
C 887	15	0.5	4060	3	US-08-845-623-3	Sequence 3, App11	C 960	15	0.5	4723	4	US-09-042-353-370	Sequence 370, App
C 888	15	0.5	4060	3	US-08-845-623-5	Sequence 3, App11	C 961	15	0.5	4723	4	US-09-042-353-370	Sequence 370, App
C 889	15	0.5	4060	3	US-08-845-623-7	Sequence 7, App11	C 962	15	0.5	4723	4	US-08-758-417A-218	Sequence 218, App
C 890	15	0.5	4060	3	US-08-845-623-9	Sequence 9, App11	C 963	15	0.5	4723	4	US-08-758-417A-218	Sequence 218, App
C 891	15	0.5	4060	3	US-08-815-927-1	Sequence 1, App11	C 964	15	0.5	4853	4	US-08-881-450A-22	Sequence 22, App1
C 892	15	0.5	4060	3	US-08-815-927-3	Sequence 3, App11	C 965	15	0.5	4853	4	US-08-881-450A-22	Sequence 22, App1
C 893	15	0.5	4060	3	US-08-815-927-5	Sequence 5, App11	C 966	15	0.5	4856	1	US-09-042-353-418	Sequence 418, App
C 894	15	0.5	4060	3	US-08-815-927-7	Sequence 7, App11	C 967	15	0.5	4856	1	US-09-042-353-418	Sequence 418, App
C 895	15	0.5	4060	4	US-09-103-330-1	Sequence 1, App11	C 968	15	0.5	4926	4	US-08-758-417A-268	Sequence 268, App
C 896	15	0.5	4060	4	US-09-103-330-3	Sequence 3, App11	C 969	15	0.5	4926	4	US-08-758-417A-268	Sequence 268, App
C 897	15	0.5	4060	4	US-09-103-330-5	Sequence 5, App11	C 970	15	0.5	4937	2	US-08-622-166A-3	Sequence 3, App11
C 898	15	0.5	4060	4	US-09-103-330-7	Sequence 7, App11	C 971	15	0.5	4990	3	US-08-776-511-1	Sequence 1, App11
C 899	15	0.5	4080	2	US-08-353-485-7	Sequence 7, App11	C 972	15	0.5	5089	6	5177197-31	Patent No. 5177197
C 900	15	0.5	4080	2	US-08-353-485-7	Sequence 7, App11	C 973	15	0.5	5097	4	US-09-315-793-11	Sequence 11, App1
C 901	15	0.5	4104	1	US-07-998-003A-94	Sequence 94, App1	C 974	15	0.5	5175	2	US-08-242-677-1	Sequence 1, App11
C 902	15	0.5	4104	1	US-08-453-274B-94	Sequence 94, App1	C 975	15	0.5	5175	2	US-08-843-530B-3	Sequence 3, App11
C 903	15	0.5	4104	1	US-08-453-695A-94	Sequence 94, App1	C 976	15	0.5	5211	1	US-08-447-411-1	Sequence 1, App11


```
977 15 0.5 5337 2 US-08-784-512-4 Sequence 4, Appl1
978 15 0.5 5337 2 US-08-784-512-4 Sequence 4, Appl1
979 15 0.5 5337 2 US-09-176-228-4 Sequence 4, Appl1
c 980 15 0.5 5337 4 US-09-176-228-4 Sequence 4, Appl1
c 981 15 0.5 5394 2 US-08-688-376-1 Sequence 1, Appl1
c 982 15 0.5 5452 2 US-09-130-114-1 Sequence 1, Appl1
c 983 15 0.5 5470 1 US-08-441-139-12 Sequence 12, Appl1
c 984 15 0.5 5470 6 5196523-5 Patent No. 5196523
985 15 0.5 5484 4 US-09-632-580A-3 Sequence 3, Appl1
986 15 0.5 5501 1 US-08-484-438-1 Sequence 1, Appl1
c 987 15 0.5 5541 1 US-08-920-812-20 Sequence 20, Appl1
c 988 15 0.5 5541 1 US-08-920-827-20 Sequence 20, Appl1
c 989 15 0.5 5541 1 US-08-921-177-20 Sequence 20, Appl1
c 990 15 0.5 5541 1 US-08-362-577C-20 Sequence 20, Appl1
c 991 15 0.5 5541 2 US-08-920-828-20 Sequence 20, Appl1
c 992 15 0.5 5655 2 US-08-989-478-1 Sequence 1, Appl1
c 993 15 0.5 5655 3 US-08-986-685-1 Sequence 1, Appl1
c 994 15 0.5 5655 3 US-08-880-179-2 Sequence 2, Appl1
995 15 0.5 5775 1 US-08-306-691B-15 Sequence 15, Appl1
996 15 0.5 5775 5 PCT-US93-06251-29 Sequence 29, Appl1
c 997 15 0.5 5822 3 US-08-899-595-4 Sequence 4, Appl1
c 998 15 0.5 5822 3 US-08-899-595-5 Sequence 5, Appl1
c 999 15 0.5 5924 1 US-08-447-411-44 Sequence 44, Appl1
c1000 15 0.5 5948 2 US-08-662-227-1 Sequence 1, Appl1
```

ALIGNMENTS

```
RESULT 1
US-09-588-256-9
; Sequence 9, Application US/09588256
; Patent No. 6291665
; GENERAL INFORMATION:
; APPLICANT: Gafney, Thomas
; APPLICANT: Gafney, Albert
; APPLICANT: Gates, Krista
; APPLICANT: Wendland, Juergen
; APPLICANT: Ayad-Durieux, Yasmina
; APPLICANT: Dietrich, Fred
; APPLICANT: Philippson, Peter
; TITLE OF INVENTION: Fungal Target Genes and Methods
; FILE REFERENCE: PB/5-30908A
; CURRENT APPLICATION NUMBER: US/09/588,256
; CURRENT FILING DATE: 2000-06-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2382
; TYPE: DNA
; ORGANISM: Ashbya gossypii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2382)
US-09-588-256-9
```

```
Query Match 0.8%; Score 21; DB 4; Length 2382;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 21; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1875 ggccttgcctaagtaact 1895
Db 1990 ggccttgcctaagtaact 2010
```

```
RESULT 2
US-08-933-750C-66
; Sequence 66, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
```

```
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1892 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTNOT03
; CLONE: 641127
US-08-933-750C-66
```

```
Query Match 0.7%; Score 20; DB 2; Length 1892;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 20; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1419 cacacagagagaacctta 1438
Db 1289 CACACAGAGAGAACCTTA 1308
```

```
RESULT 3
US-09-234-613-66
; Sequence 66, Application US/09234613
; Patent No. 6139973
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
```


STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/234,613
 FILING DATE:
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US/08/933,750
 FILING DATE: September 23, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0356 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 TELEX:
 INFORMATION FOR SEQ ID NO: 66:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1892 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: BRESTNOT03
 CLONE: 641127
 US-09-234-613-66

Query Match 0.7%; Score 20; DB 3; Length 1892;
 Best Local Similarity 100.0%; Pred. No. 3.4;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1419 cacacagagagagacccta 1438
 ||||||||||||||||||
 DB 1289 CACACAGAGAGAGACCTTA 1308

RESULT 4
 US-09-262-773-7
 ; Sequence 7, Application US/09262773
 ; Patent No. 6225451
 ; GENERAL INFORMATION:
 ; APPLICANT: Ballinger, Dennis G.
 ; APPLICANT: Ding, Wei
 ; APPLICANT: Wagner, Susanne
 ; APPLICANT: Hess, Mark A.
 ; TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
 ; FILE REFERENCE: Myriad 3
 ; CURRENT APPLICATION NUMBER: US/09/262,773
 ; CURRENT FILING DATE: 1999-03-04
 ; NUMBER OF SEQ ID NOS: 210
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 7
 ; LENGTH: 3240
 ; TYPE: DNA
 ; ORGANISM: human
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (98)..(523)
 ; US-09-262-773-7

Query Match 0.7%; Score 19; DB 4; Length 3240;

Best Local Similarity 100.0%; Pred. No. 11;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2068 tcagacatcagagacaca 2086
 ||||||||||||||||||
 DB 1559 tcagacatcagagacaca 1577

RESULT 5
 US-09-262-773-3
 ; Sequence 3, Application US/09262773
 ; Patent No. 6225451
 ; GENERAL INFORMATION:
 ; APPLICANT: Ballinger, Dennis G.
 ; APPLICANT: Ding, Wei
 ; APPLICANT: Wagner, Susanne
 ; APPLICANT: Hess, Mark A.
 ; TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
 ; FILE REFERENCE: Myriad 3
 ; CURRENT APPLICATION NUMBER: US/09/262,773
 ; CURRENT FILING DATE: 1999-03-04
 ; NUMBER OF SEQ ID NOS: 210
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 3244
 ; TYPE: DNA
 ; ORGANISM: human
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (98)..(2017)
 ; US-09-262-773-3

Query Match 0.7%; Score 19; DB 4; Length 3244;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2068 tcagacatcagagacaca 2086
 ||||||||||||||||||
 DB 1563 tcagacatcagagacaca 1581

RESULT 6
 US-09-262-773-5
 ; Sequence 5, Application US/09262773
 ; Patent No. 6225451
 ; GENERAL INFORMATION:
 ; APPLICANT: Ballinger, Dennis G.
 ; APPLICANT: Ding, Wei
 ; APPLICANT: Wagner, Susanne
 ; APPLICANT: Hess, Mark A.
 ; TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
 ; FILE REFERENCE: Myriad 3
 ; CURRENT APPLICATION NUMBER: US/09/262,773
 ; CURRENT FILING DATE: 1999-03-04
 ; NUMBER OF SEQ ID NOS: 210
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 3264
 ; TYPE: DNA
 ; ORGANISM: human
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (98)..(496)
 ; US-09-262-773-5

Query Match 0.7%; Score 19; DB 4; Length 3264;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2068 tcagacatcagagacaca 2086
Db 1583 tcagacatcagagacaca 1601

RESULT 7

US-09-262-773-1
; Sequence 1, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; APPLICANT: Ding, Wei
; APPLICANT: Wagner, Susanne
; APPLICANT: Hess, Mark A.
; TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
; FILE REFERENCE: Myriad 3
; CURRENT APPLICATION NUMBER: US/09/262,773
; CURRENT FILING DATE: 1999-03-04
; NUMBER OF SEQ ID NOS: 210
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3268
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (98)..(2041)
US-09-262-773-1

Query Match 0.7%; Score 19; DB 4; Length 3268;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY : 2068 tcagacatcagagacaca 2086
Db 1587 tcagacatcagagacaca 1605

RESULT 8

US-08-475-844-8
; Sequence 8, Application US/08475844
; Patent No. 5972643
; GENERAL INFORMATION:
; APPLICANT: Lobanenko, Victor V.
; APPLICANT: Neiman, Paul E.
; APPLICANT: Klenova, Elena M.
; APPLICANT: Goodwin, Graham H.
; APPLICANT: Filippova, Galina N.
; APPLICANT: Collins, Steven J.
; TITLE OF INVENTION: CTCF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,844
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/261,660
; FILING DATE: 17-JUN-1994
; CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14538A-11-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-543-5043

; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3810 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: human CTCF cDNA

; FEATURE:
; NAME/KEY: CDS
; LOCATION: 292..2475
; FEATURE:
; NAME/KEY: exon
; LOCATION: 281..1074
; OTHER INFORMATION: /label= exon2
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1075..1245
; OTHER INFORMATION: /label= exon3
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1246..1379
; OTHER INFORMATION: /label= exon4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1380..1499
; OTHER INFORMATION: /label= exon5
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1500..1649
; OTHER INFORMATION: /label= exon6
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1650..1810
; OTHER INFORMATION: /label= exon7
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1810..1992
; OTHER INFORMATION: /label= exon8
US-08-475-844-8

Query Match 0.7%; Score 19; DB 2; Length 3810;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1755 catcaggggaaaagcctt 1773
Db 1492 catcaggggaaaagcctt 1510

RESULT 9

PCT-US95-08429-8
; Sequence 8, Application PC/TUS9508429
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: CTCF
; NUMBER OF SEQUENCES: 21
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08429
FILING DATE: 15-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/261,680
FILING DATE: 17-JUN-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14538A-11-1PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-467-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3810 base pairs
TYPE: nucleic acid
STRADEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: human CTCF cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 292..2475
FEATURE:
NAME/KEY: exon
LOCATION: 281..1074
OTHER INFORMATION: /label= exon2
FEATURE:
NAME/KEY: exon
LOCATION: 1075..1245
OTHER INFORMATION: /label= exon3
FEATURE:
NAME/KEY: exon
LOCATION: 1246..1379
OTHER INFORMATION: /label= exon4
FEATURE:
NAME/KEY: exon
LOCATION: 1380..1499
OTHER INFORMATION: /label= exon5
FEATURE:
NAME/KEY: exon
LOCATION: 1500..1649
OTHER INFORMATION: /label= exon6
FEATURE:
NAME/KEY: exon
LOCATION: 1650..1810
OTHER INFORMATION: /label= exon7
FEATURE:
NAME/KEY: exon
LOCATION: 1810..1992
OTHER INFORMATION: /label= exon8
PCT-US95-08429-8

```

```

Query Match          0.7%; Score 19; DB 5; Length 3810;
Best Local Similarity 100.0%; Pred. NO. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1755 catcaggagaaagccct 1773
      |||||
Db 1492 CATTCAGGGGAAAGCCTT 1510

RESULT 10
US-09-262-773-206
; Sequence 206, Application US/09262773
; Patent No. 6225451

```

```

1  GENERAL INFORMATION:
2  APPLICANT: Ballinger, Dennis G.
3  APPLICANT: Ding, Wei
4  APPLICANT: Wagner, Susanne
5  APPLICANT: Hess, Mark A.
6  TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
7  TITLE OF INVENTION: SUSCEPTIBILITY GENE CHD1
8  FILE REFERENCE: Myriad 3
9  CURRENT APPLICATION NUMBER: US/09/262,773
10 CURRENT FILING DATE: 1999-03-04
11 NUMBER OF SEQ ID NOS: 210
12 SOFTWARE: PatentIn Ver. 2.0
13 SEQ ID NO 206
14 LENGTH: 20137
15 TYPE: DNA
16 ORGANISM: human
17 US-09-262-773-206

```

Query Match	0.7%	Score 19	DB 4	Length 20137
Best Local Similarity	100.0%	Pred. No. 10		
Matches 19	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 2068	tcagacatcagagagacaca	2086		
Db 14680	tcagacatcagagagacaca	14698		

```

RESULT 11
US-09-262-773-9
: Sequence 9, Application US/09262773
: Patent No. 6225451
: GENERAL INFORMATION:
: APPLICANT: Ballinger, Dennis G.
: APPLICANT: Ding, Wei
: APPLICANT: Wagner, Susanne
: APPLICANT: Hess, Mark A.
: TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
: TITLE OF INVENTION: SUSCEPTIBILITY GENE CHD1
: FILE REFERENCE: Myriad 3
: CURRENT APPLICATION NUMBER: US/09/262,773
: CURRENT FILING DATE: 1999-03-04
: NUMBER OF SEQ ID NOS: 210
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 9
: LENGTH: 20138
: TYPE: DNA
: ORGANISM: human
: US-09-262-773-9

```

	Query Match	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps
	0.7%	Score 19; DB 4;	length 20138;				
		100.0%;	Pred. No. 10;				
				0;	0;	0;	0;
QY	2068	tcagacatcagagacaca	2086				
db	14681	tcagacatcagagacaca	14699				

RESULT 12
US-09-262-773-210
; Sequence 210, Application US/09262773
; Patent No. 6225451
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Dennis G.
; APPLICANT: ding, wei
; APPLICANT: wagner, Susanne
; APPLICANT: Hess, Mark A.
; TITLE OF INVENTION: CHROMOSOME 11-LINKED CORONARY HEART DISEASE
; TITLE OF INVENTION: SUSCEPTIBILITY GENE CHD1
; FILE REFERENCE: Myriad 3
; CURRENT APPLICATION NUMBER: US/09/262.773

;; CURRENT FILING DATE: 1999-03-04
;; NUMBER OF SEQ ID NOS: 210
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 210
;; LENGTH: 23071
;; TYPE: DNA
;; ORGANISM: human
US-09-262-773-210

Query Match 0.7%; Score 19; DB 4; Length 23071;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2068 tcagacatcagagacaca 2086
|||||
Db 17614 tcagacatcagagacaca 17632

RESULT 13
US-08-724-394A-20/C

; Sequence 20, Application US/08724394A
; Patent No. 5872237

;; GENERAL INFORMATION:

;; APPLICANT: Feder, John N.

;; APPLICANT: Krommal, Gregory S.

;; APPLICANT: Laufer, Peter M.

;; APPLICANT: Ruddy, David A.

;; APPLICANT: Thomas, Winston

;; APPLICANT: Tsuchihashi, Zenta

;; APPLICANT: Wolff, Roger K.

;; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1

;; TITLE OF INVENTION: Sequences and Antibodies Thereo

;; NUMBER OF SEQUENCES: 31

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: TOWNSEND and CREW LLP

;; STREET: Two Embarcadero Center, 8th Floor

;; CITY: San Francisco

;; STATE: CA

;; COUNTRY: USA

;; ZIP: 94111-3834

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/724,394A

;; FILING DATE: 01-OCT-1996

;; CLASSIFICATION: 536

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Filts, Renee A.

;; REGISTRATION NUMBER: 35,136

;; REFERENCE/DOCKET NUMBER: 017957-000100

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: 415-576-0200

;; TELEFAX: 415-576-0300

;; INFORMATION FOR SEQ ID NO: 20:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 246240 base pairs

;; TYPE: nucleic acid

;; STRANDEDNESS: not relevant

;; TOPOLOGY: not relevant

;; MOLECULE TYPE: cDNA

;; FEATURE:

;; NAME/KEY: misc_feature

;; LOCATION: 1..246240

;; OTHER INFORMATION: /note= "HLA-H-CONTIG"

US-08-724-394A-20

Query Match 0.7%; Score 19; DB 2; Length 246240;

Best Local Similarity 100.0%; Pred. No. 10;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2298 tttagagaccctgcctt 2316
|||||
Db 36115 TTTCAGAGACCCCTGCCTT 36097

RESULT 14

US-08-724-394A-21/C

; Sequence 21, Application US/08724394A

; Patent No. 5872237

;; GENERAL INFORMATION:

;; APPLICANT: Feder, John N.

;; APPLICANT: Krommal, Gregory S.

;; APPLICANT: Laufer, Peter M.

;; APPLICANT: Ruddy, David A.

;; APPLICANT: Thomas, Winston

;; APPLICANT: Tsuchihashi, Zenta

;; APPLICANT: Wolff, Roger K.

;; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1

;; TITLE OF INVENTION: Sequences and Antibodies Thereo

;; NUMBER OF SEQUENCES: 31

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: TOWNSEND and CREW LLP

;; STREET: Two Embarcadero Center, 8th Floor

;; CITY: San Francisco

;; STATE: CA

;; COUNTRY: USA

;; ZIP: 94111-3834

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/724,394A

;; FILING DATE: 01-OCT-1996

;; CLASSIFICATION: 536

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Filts, Renee A.

;; REGISTRATION NUMBER: 35,136

;; REFERENCE/DOCKET NUMBER: 017957-000100

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: 415-576-0200

;; TELEFAX: 415-576-0300

;; INFORMATION FOR SEQ ID NO: 21:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 246240 base pairs

;; TYPE: nucleic acid

;; STRANDEDNESS: not relevant

;; TOPOLOGY: not relevant

;; MOLECULE TYPE: cDNA

;; FEATURE:

;; NAME/KEY: misc_feature

;; LOCATION: 1..246240

;; OTHER INFORMATION: /note= "HLA-H-CONTIG"

US-08-724-394A-21

Query Match 0.7%; Score 19; DB 2; Length 246240;

Best Local Similarity 100.0%; Pred. No. 10;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2298 tttagagaccctgcctt 2316
|||||

Db 36115 TTTCAGAGACCCCTGCCTT 36097

RESULT 15

US-08-724-394A-22/C

; Sequence 22, Application US/08724394A

; Patent No. 5872237

;; GENERAL INFORMATION:

APPLICANT: Feder, John N.
APPLICANT: Kironmal, Gregory S.
APPLICANT: Lauer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
TITLE OF INVENTION: Sequences and Antibodies Thereeto
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 01/957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ. ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc-feature
LOCATION: 1..246240
OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-22

Query Match 0.7%; Score 19; DB 2; Length 246240;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2298 ttccagagaccctgcctt 2316
Db 36115 ttccagagaccctgcctt 36097

RESULT 16
US-08-507-016-8
Sequence 8, Application US/08507016
Patent No. 5736460
GENERAL INFORMATION:
APPLICANT: EVANS, HELEN F.
APPLICANT: SHINE, JOHN
TITLE OF INVENTION: HUMAN GALANIN, cDNA CLONES ENCODING
TITLE OF INVENTION: HUMAN GALANIN AND A METHOD OF PRODUCING HUMAN GALANIN
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
STREET: 555 THIRTEENTH STREET, N.W.
CITY: WASHINGTON
STATE: D. C.
COUNTRY: USA
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/507,016
FILING DATE: 25-JULY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/108,733
FILING DATE: 03-SEP-1993
APPLICATION NUMBER: PCT/AU92/00097
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: AU PK4953
FILING DATE: 06-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: ERNST, BARBARA G.
REGISTRATION NUMBER: 30,377
REFERENCE/DOCKET NUMBER: 1871-117A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
TELEFAX: (202)783-6031
INFORMATION FOR SEQ. ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 575 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 14..385
US-08-507-016-8

Query Match 0.6%; Score 18; DB 1; Length 575;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1553 cctgaagtcacacctaa 1570
Db 420 cctgaagtcacaccttaa 437

RESULT 17
US-09-303-524A-1/c
Sequence 1, Application US/09303524A
Patent No. 6238873
GENERAL INFORMATION:
APPLICANT:
APPLICANT: CHAMBERS, JONATHAN K.
APPLICANT: STEWART, BRIAN R.
APPLICANT: AMES, ROBERT S.
APPLICANT: SARAU, HENRY M.
APPLICANT: FOLEY, JIM
APPLICANT: ARNOLD, ANNE ROMANIC
TITLE OF INVENTION: METHODS OF SCREENING FOR AGONISTS AND
TITLE OF INVENTION: ANTAGONISTS OF THE INTERACTION BETWEEN THE HUMAN KIAA0001
FILE REFERENCE: GP50007
CURRENT APPLICATION NUMBER: US/09/303,524A
CURRENT FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: 60/083,957
PRIOR FILING DATE: 1998-05-01
NUMBER OF SEQ. ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ. ID NO 1
LENGTH: 1027
TYPE: DNA
ORGANISM: Homo sapiens
US-09-303-524A-1

Query Match 0.6%; Score 18; DB 4; Length 1027;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 570 aaacatcatccagaagat 587
|||||
DB 600 AAACATCATCCAGAAGAT 583

RESULT 18

US-08-416-870C-9
; Sequence 9, Application US/08416870C
; Patent No. 5824862
; GENERAL INFORMATION:
; APPLICANT: HIYOSHI, TORU
; APPLICANT: MINE, TOSHIKI
; APPLICANT: KASAOKA, KEISUKE
; APPLICANT: TYSON, ROBERT HUM
; APPLICANT: PAGE, ANTHONY MILES JOHN
; TITLE OF INVENTION: DNA ENCODING ATP-DEPENDENT FRUCTOSE 6-PHOSPHATE
; TITLE OF INVENTION: 1-PHOSPHOTRANSFERASE ORIGINATING FROM PLANT, RECOMBINANT VECTOR
; TITLE OF INVENTION: SAME AND METHOD FOR CHANGING SUGAR CONTENT IN PLANT CELLS UNDER
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
; STREET: PO BOX 747
; CITY: FALL CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,870C
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MORPHY JR, GERALD M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 760-195P(PCT)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1558 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1278
; US-08-416-870C-9

Query Match 0.6%; Score 18; DB 1; Length 1558;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1866 tctgggcaagccttctgt 1883
|||||
DB 1175 TGTGGCAAGCCTTTGCT 1192

RESULT 19

US-08-820-170A-11
; Sequence 11, Application US/08820170A

; Patent No. 5831058
; GENERAL INFORMATION:
; APPLICANT: TSUTOMU, FUJIWARA
; APPLICANT: TAKESHI, WATANABE
; APPLICANT: MASATO, HORIE
; APPLICANT: TOYOMASA, KATAGIRI
; TITLE OF INVENTION: HUMAN GENE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/820,170A
; FILING DATE:
; CLASSIFICATION: 536
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2133 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA(genomic)
; US-08-820-170A-11

Query Match 0.6%; Score 18; DB 2; Length 2133;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1087 caaggagaaacctatg 1104
|||||
DB 1493 CAGGGAGAAACCTTATG 1510

RESULT 20

US-09-055-699-11
; Sequence 11, Application US/09055699
; Patent No. 6005088
; GENERAL INFORMATION:
; APPLICANT: TSUTOMU, FUJIWARA
; APPLICANT: TAKESHI, WATANABE
; APPLICANT: MASATO, HORIE
; APPLICANT: TOYOMASA, KATAGIRI
; TITLE OF INVENTION: HUMAN GENE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/055,699
; FILING DATE:

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/820,170
; FILING DATE:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2133 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA(genomic)
; US-09-055-699-11

Query Match          0.6%; Score 18; DB 3; Length 2133;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1087 caggggagaacctatg 1104
      |||
Db 1493 CAGGGGAGAAACCTATG 1510

RESULT 21
US-09-273-565-11
; Sequence 11, Application US/09273565A
; Patent No. 6166190
; GENERAL INFORMATION:
; APPLICANT: FUJIMURA, TSUTOMU
; APPLICANT: WATANABE, TAKESHI
; APPLICANT: HORIE, MASATO
; TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
; TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
; FILE REFERENCE: Q-53599
; CURRENT APPLICATION NUMBER: US/09/273,565A
; CURRENT FILING DATE: 1999-03-22
; EARLIER APPLICATION NUMBER: 08/055,699
; EARLIER FILING DATE: 1998-04-07
; EARLIER APPLICATION NUMBER: 08/820,170
; EARLIER FILING DATE: 1997-03-19
; EARLIER APPLICATION NUMBER: JP 63410/1996
; EARLIER FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: JP 69163/1997
; EARLIER FILING DATE: 1997-03-05
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 2133
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-273-565-11

Query Match          0.6%; Score 18; DB 4; Length 2133;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1087 caggggagaacctatg 1104
      |||
Db 1493 caggggagaacctatg 1510

RESULT 22
US-09-565-538-11
; Sequence 11, Application US/09565538
; Patent No. 6333404
; GENERAL INFORMATION:
; APPLICANT: FUJIMURA, TSUTOMU
; APPLICANT: WATANABE, TAKESHI
```

```

; APPLICANT: HORIE, MASATO
; TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
; TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
; FILE REFERENCE: Q-53599
; CURRENT APPLICATION NUMBER: US/09/565,538
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: 09/273,565
; PRIOR FILING DATE: 1999-03-22
; PRIOR APPLICATION NUMBER: 09/055,699
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: 08/820,170
; PRIOR FILING DATE: 1997-03-19
; PRIOR APPLICATION NUMBER: JP 63410/1996
; PRIOR FILING DATE: 1996-03-19
; PRIOR APPLICATION NUMBER: JP 69163/1997
; PRIOR FILING DATE: 1997-03-05
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 2133
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-565-538-11

Query Match          0.6%; Score 18; DB 4; Length 2133;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1087 caggggagaacctatg 1104
      |||
Db 1493 caggggagaacctatg 1510

RESULT 23
US-08-820-170A-12
; Sequence 12, Application US/08820170A
; Patent No. 5831058
; GENERAL INFORMATION:
; APPLICANT: Tsutomu, FUJIMURA
; APPLICANT: Takeshi, WATANABE
; APPLICANT: Masato, HORIE
; APPLICANT: Toyomasa, KATAGIRI
; TITLE OF INVENTION: HUMAN GENE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zimm, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/820,170A
; FILING DATE:
; CLASSIFICATION: 536
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3754 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA(genomic)
; HYPOTHETICAL: NO
```

ANTI-SENSE: NO
IMMEDIATE SOURCE:
LIBRARY: Human fetal brain cDNA library
CLONE: GEN-076C09
FEATURE:
NAME/KEY: CDS
LOCATION: 346..2478
US-08-820-170A-12

Query Match 0.6%; Score 18; DB 2; Length 3754;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1087 caggggagaaacctatg 1104
|||||
DB 1838 CAGGGGAGAAACCTTATG 1855

RESULT 24
US-09-055-699-12
Sequence 12, Application US/09055699
Patent No. 6005088
GENERAL INFORMATION:
APPLICANT: Tsutomu, FUJIWARA
APPLICANT: Takeshi, WATANABE
APPLICANT: Masato, HORIE
APPLICANT: Toyomasa, KATAGIRI
TITLE OF INVENTION: HUMAN GENE
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/055,699
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/820,170
FILING DATE:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 293-7060
TELEFAX: (202) 293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 3754 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA(genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
LIBRARY: Human fetal brain cDNA library
CLONE: GEN-076C09
FEATURE:
NAME/KEY: CDS
LOCATION: 346..2478
US-09-055-699-12

Query Match 0.6%; Score 18; DB 3; Length 3754;
Best Local Similarity 100.0%; Pred. No. 32;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1087 caggggagaaacctatg 1104
|||||
DB 1838 CAGGGGAGAAACCTTATG 1855

RESULT 25
US-09-273-565-12
Sequence 12, Application US/09273565A
Patent No. 6166190
GENERAL INFORMATION:
APPLICANT: FUJIWARA, TSUTOMU
APPLICANT: WATANABE, TAKESHI
APPLICANT: HORIE, MASATO
TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
FILE REFERENCE: Q-53599
CURRENT APPLICATION NUMBER: US/09/273,565A
CURRENT FILING DATE: 1999-03-22
EARLIER APPLICATION NUMBER: 09/055,699
EARLIER FILING DATE: 1998-04-07
EARLIER APPLICATION NUMBER: 08/820,170
EARLIER FILING DATE: 1997-03-19
EARLIER APPLICATION NUMBER: JP 63410/1996
EARLIER FILING DATE: 1996-03-19
EARLIER APPLICATION NUMBER: JP 69163/1997
EARLIER FILING DATE: 1997-03-05
NUMBER OF SEQ ID NOS: 95
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 3754
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (346)..(2478)
US-09-273-565-12

Query Match 0.6%; Score 18; DB 4; Length 3754;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1087 caggggagaaacctatg 1104
|||||
DB 1838 caggggagaaacctatg 1855

RESULT 26
US-09-565-538-12
Sequence 12, Application US/09565538
Patent No. 633404
GENERAL INFORMATION:
APPLICANT: FUJIWARA, TSUTOMU
APPLICANT: WATANABE, TAKESHI
APPLICANT: HORIE, MASATO
TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
FILE REFERENCE: Q-53599
CURRENT APPLICATION NUMBER: US/09/565,538
CURRENT FILING DATE: 2000-05-05
PRIOR APPLICATION NUMBER: 09/273,565
PRIOR FILING DATE: 1999-03-22
PRIOR APPLICATION NUMBER: 09/055,699
PRIOR FILING DATE: 1998-04-07
PRIOR APPLICATION NUMBER: 08/820,170
PRIOR FILING DATE: 1997-03-19
PRIOR APPLICATION NUMBER: JP 63410/1996
PRIOR FILING DATE: 1996-03-19
PRIOR APPLICATION NUMBER: JP 69163/1997
PRIOR FILING DATE: 1997-03-05
NUMBER OF SEQ ID NOS: 95


```
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 12
; LENGTH: 3754
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (346)..(2478)
; US-09-565-538-12
```

```
Query Match          0.6%; Score 18; DB 4; Length 3754;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1087 caggggagaacattatg 1104
      |||||||
Db 1838 caggggagaacattatg 1855
```

RESULT 27

```
US-08-751-189-2/C
; Sequence 2, Application US/08751189
; Patent No. 5919656
```

```
; GENERAL INFORMATION:
; APPLICANT: Harrington, Lea A.
; APPLICANT: Robinson, Murray O.
; TITLE OF INVENTION: No. 5919656e1 Genes Encoding Telomerase Protein
; NUMBER OF SEQUENCES: 12
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen, Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
```

```
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/751,189
; FILING DATE: 15-NOV-1996
```

```
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-433
```

```
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7886 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: cDNA
; US-08-751-189-2
```

```
Query Match          0.6%; Score 18; DB 2; Length 7886;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 2553 ttttactcgtctgtgc 2570
      |||||||
Db 7660 tttttactcgtctgtgc 7643
```

```
RESULT 28
US-09-060-836-2/C
; Sequence 2, Application US/09060836
; Patent No. 5981707
```

```
; GENERAL INFORMATION:
; APPLICANT: Harrington, Lea A.
; APPLICANT: Robinson, Murray O.
; TITLE OF INVENTION: No. 5981707e1 Genes Encoding Telomerase Protein
; NUMBER OF SEQUENCES: 12
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen, Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
```

```
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,836
; FILING DATE:
```

```
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/751,189
; FILING DATE: 15-NOV-1996
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy A.
; REGISTRATION NUMBER: 34,688
; REFERENCE/DOCKET NUMBER: A-433
```

```
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7886 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: cDNA
; US-09-060-836-2
```

```
Query Match          0.6%; Score 18; DB 2; Length 7886;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 2553 ttttactcgtctgtgc 2570
      |||||||
Db 7660 tttttactcgtctgtgc 7643
```

```
RESULT 29
US-09-184-445-2/C
; Sequence 2, Application US/09184445
; Patent No. 6174703
```

```
; GENERAL INFORMATION:
; APPLICANT: Harrington, Lea A.
; APPLICANT: Robinson, Murray O.
; TITLE OF INVENTION: No. 6174703e1 Genes Encoding Telomerase Protein
; NUMBER OF SEQUENCES: 12
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen, Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
```

```
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/184,445
; FILING DATE:
```

```
US-09-184-445-2/C
; Sequence 2, Application US/09184445
; Patent No. 6174703
```

```
; GENERAL INFORMATION:
; APPLICANT: Harrington, Lea A.
; APPLICANT: Robinson, Murray O.
; TITLE OF INVENTION: No. 6174703e1 Genes Encoding Telomerase Protein
; NUMBER OF SEQUENCES: 12
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen, Inc.
; STREET: 1840 De Havilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/184,445
; FILING DATE:
```

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/751,189
FILING DATE: 15-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Oleski, Nancy A.
REGISTRATION NUMBER: 34,688
REFERENCE/DOCKET NUMBER: A-433
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 7886 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-184-445-2

Query Match 0.6%; Score 18; DB 4; Length 7886;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2553 ttttacctgcgtgcc 2570
|||||
Db 7660 tttttacctgcgtgcc 7643

RESULT 30
PCT-US91-02942-99
Sequence 99, Application PC/TUS9102942
GENERAL INFORMATION:
APPLICANT: ROTHLEIN, ROBERT
APPLICANT: ADAIR, JOHN R.
APPLICANT: ATHAL, DILEET S.
TITLE OF INVENTION: HUMANIZED CDR-GRAFTED ICAM-1 ANTIBODY
NUMBER OF SEQUENCES: 102
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Ave. NW Suite 300
City: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/02942
FILING DATE: 19910429
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9009549.8
FILING DATE: 27-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: FOX, SAM L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 1011.0586600
TELEPHONE: (202) 466-0800
TELEFAX: (202) 833-8716
INFORMATION FOR SEQ ID NO: 99:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US91-02942-99

Query Match 0.6%; Score 17; DB 5; Length 80;

Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 650 agactccagactcctgt 666
|||||
Db 11 AGACTCCAGACTCCTGT 27

RESULT 31
US-08-956-182-21/c
Sequence 21, Application US/08956182
Patent No. 6100450
GENERAL INFORMATION:
APPLICANT: Thomas, Terry L.
TITLE OF INVENTION: NOVEL SEED SPECIFIC PROMOTERS BASED ON
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
City: Garden City
STATE: New York
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,182
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: DIGILIO, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 10701
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 273 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-956-182-21

Query Match 0.6%; Score 17; DB 3; Length 273;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2101 ttatgtgtataggaa 2117
|||||
Db 213 TTTATGTGTATAGGAA 197

RESULT 32
US-08-592-126-124/c
Sequence 124, Application US/08592126
Patent No. 5821091
GENERAL INFORMATION:
APPLICANT: Gregory Dolganov
TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
NUMBER OF SEQUENCES: 151
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
City: Palo Alto

STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,126
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 4600-0111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 124:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: G45.seq
US-08-592-126-124

Query Match 0.6%; Score 17; DB 1; Length 289;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 390 tggagagagagagaaa 406
|||||
DB 262 TGGAGAGAGAGAGAAA 246

RESULT 33
US-08-905-223-133
Sequence 133, Application US/08905223
Patent No. 6222029
GENERAL INFORMATION:
APPLICANT: Edwards, Jean-Baptiste D.
APPLICANT: Duclert, Aymeric
APPLICANT: Lacroix, Bruno
TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
NUMBER OF SEQUENCES: 503
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knudde, Martens, Olsson & Bear
STREET: 501 West Broadway
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-3505
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/905,223
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 133:
SEQUENCE CHARACTERISTICS:
LENGTH: 347 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 183..338
IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 4.1
OTHER INFORMATION: seq vmlctgclvslg/hp
US-08-905-223-133

Query Match 0.6%; Score 17; DB 4; Length 347;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 294 gaggtgatgctgagagac 310
|||||
DB 297 GAGGTGATGCTGAGAGAC 313

RESULT 34
US-09-328-111-648
Sequence 648, Application US/09328111
Patent No. 6262333
GENERAL INFORMATION:
APPLICANT: Endege, Wilson O.
APPLICANT: Steinmann, Kathleen E.
APPLICANT: Astle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Catino, Theodore J.
APPLICANT: Dertl, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/328,111
CURRENT FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: US 60/088,801
EARLIER FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 648
LENGTH: 532
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(532)
OTHER INFORMATION: n = A,T,C or G
US-09-328-111-648

Query Match 0.6%; Score 17; DB 4; Length 532;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1251 cacactgggagagagcc 1267
|||||

Db 254 cacactggtgaggaagcc 270

RESULT 35

US-09-385-982-433/C
; Sequence 433, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS: II
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 433
; LENGTH: 563
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(563)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-433

Query Match 0.6%; Score 17; DB 4; Length 563;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1419 cacacaggaagaagcc 1435
|||
Db 145 CACACAGGAGAGGAGGCC 129

RESULT 36

US-08-956-182-16/C
; Sequence 16, Application US/08956182
; Patent No. 6100450
; GENERAL INFORMATION:
; APPLICANT: Thomas, Terry L.
; TITLE OF INVENTION: NOVEL SEED SPECIFIC PROMOTERS BASED ON
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,182
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 10701
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343

; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 732 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-956-182-16

Query Match 0.6%; Score 17; DB 3; Length 732;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2101 ttatgtgtatagga 2117
|||
Db 398 TTTATGTGTATAGGAA 382

RESULT 37

US-09-171-461-47
; Sequence 47, Application US/09171461
; Patent No. 6335016
; GENERAL INFORMATION:
; APPLICANT: Baker, Adam
; APPLICANT: Cotten, Matthew
; APPLICANT: Chiocca, Susanna
; APPLICANT: Kurzbauer, Robert
; APPLICANT: Schaffner, Gotthold
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) VIRUS
; FILE REFERENCE: 0652,1800000
; CURRENT APPLICATION NUMBER: US/09/171,461
; CURRENT FILING DATE: 1999-01-12
; EARLIER APPLICATION NUMBER: PCT/EP97/01944
; EARLIER FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47
; LENGTH: 1362
; TYPE: DNA
; ORGANISM: CELO VIRUS
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1362)
US-09-171-461-47

Query Match 0.6%; Score 17; DB 4; Length 1362;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1028 ggaatgtgagcgagct 1044
|||
Db 746 ggaatgtgagcgagct 762

RESULT 38

US-07-662-007B-38/C
; Sequence 38, Application US/07662007B
; Patent No. 5344771
; GENERAL INFORMATION:
; APPLICANT: Davies, Huw Maelor
; APPLICANT: Pollard, Michael Roman
; APPLICANT: Voelker, Toni Alois
; APPLICANT: Thompson, Gregory A.
; TITLE OF INVENTION: Plant Thioesterases
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis

STATE: California
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 KB storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.7
SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/662,007B
FILING DATE: 19910408
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/514,030
FILING DATE: 26-APR-1990
APPLICATION NUMBER: 07/620,426
FILING DATE: 30-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
REFERENCE/DOCKET NUMBER: CGNE 70-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 916-753-6313
TELEFAX: 916-753-1510
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-07-662-007B-38

Query Match 0.6%; Score 17; DB 1; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0;
Indels 0; Gaps 0;
OY 883 agagatcaactcctt 899
|||||
DB 969 AGGAGTCAAACTCCTT 953

RESULT 39
US-07-968-971A-1/c
Sequence 1, Application US/07968971A
Patent No. 545167
GENERAL INFORMATION:
APPLICANT: Toni A. Voelker
APPLICANT: Huw Maelor Davies
TITLE OF INVENTION: Medium-Chain Thioesterases
TITLE OF INVENTION: In Plants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.7
SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/968,971A
FILING DATE: 19921030
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/704,861
FILING DATE: 21-MAY-1991
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/773,096
FILING DATE: 7-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/782,263
FILING DATE: 24-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/824,247
FILING DATE: 22-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/04332
FILING DATE: 21-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
REFERENCE/DOCKET NUMBER: CGNE 88
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-07-968-971A-1

Query Match 0.6%; Score 17; DB 1; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0;
Indels 0; Gaps 0;
OY 883 agagatcaactcctt 899
|||||
DB 969 AGGAGTCAAACTCCTT 953

RESULT 40
US-07-824-247-38/c
Sequence 38, Application US/07824247
Patent No. 5512482
GENERAL INFORMATION:
APPLICANT: Voelker, Toni Alois
TITLE OF INVENTION: Plant Thioesterases
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: California
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 KB storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.7
SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/824,247
FILING DATE: 19920122
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/782,263
FILING DATE: 24-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/773,096
FILING DATE: 7-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/704,861
FILING DATE: 21-MAY-1991
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/02960
FILING DATE: 25-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/662,007
FILING DATE: 27-FEB-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/620,426
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/514,030
FILING DATE: 26-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
REFERENCE/DOCKET NUMBER: CGNE 82-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 916-753-1510
TELEFAX: 916-753-6313
INFORMATION FOR SEQ. ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-07-824-247-38

Query Match 0.6%; Score 17; DB 1; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 883 aggaagtcacacccctctt 899
|||||
DB 969 AGGAGTCAACCTCCTT 953

RESULT 41
US-07-824-247-41/C
Sequence 41, Application US/07824247
Patent No. 5512482
GENERAL INFORMATION:
APPLICANT: Voelker, Toni Alois
TITLE OF INVENTION: Plant Thioesterases
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: California
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 KB storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.7
SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/824,247
FILING DATE: 19920122
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/782,263
FILING DATE: 24-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/773,096
FILING DATE: 7-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/704,861
FILING DATE: 21-MAY-1991
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/02960
FILING DATE: 25-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/662,007
FILING DATE: 27-FEB-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/620,426
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/514,030
FILING DATE: 26-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
REFERENCE/DOCKET NUMBER: CGNE 82-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 916-753-1510
TELEFAX: 916-753-6313
INFORMATION FOR SEQ. ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-07-824-247-41

Query Match 0.6%; Score 17; DB 1; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 883 aggaagtcacacccctctt 899
|||||
DB 969 AGGAGTCAACCTCCTT 953

RESULT 42
US-08-142-473A-1/C
Sequence 1, Application US/08142473A
Patent No. 5639790
GENERAL INFORMATION:
APPLICANT: Voelker, Toni Alois
APPLICANT: Davies, Huw Maeior
TITLE OF INVENTION: Plant Thioesterases
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: California
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 KB storage
COMPUTER: Apple Macintosh IIfx
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/142,473A
FILING DATE: 18-NOV-1993
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/04332
FILING DATE: 21-MAY-1992
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/824,247
FILING DATE: 22-JAN-1992
CLASSIFICATION: 800
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/782,263
FILING DATE: 24-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/773,096
FILING DATE: 7-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/704,861
FILING DATE: 21-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 82-4 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 916-753-6313
TELEFAX: 916-753-1510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-142-473A-1

Query Match 0.6%; Score 17; DB 1; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 883 aggaagtcacacccctt 899
|||||

Db 969 AGGAGTCAACCTCCTT 953

RESULT 43

US-08-424-406-1/C
Sequence 1, Application US/08424406
Patent No. 5667997
GENERAL INFORMATION:
APPLICANT: Voelker, Toni Alois
APPLICANT: Davies, Huw Maelor
APPLICANT: Knitzon, Deborah S.
TITLE OF INVENTION: Medium Chain Thioesterases in Plants
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1(a)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424,406
FILING DATE: 26 April, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/10814
FILING DATE: 29 October, 1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 07/968,971
FILING DATE: 30 October, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen

REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 88-1US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-424-406-1

Query Match 0.6%; Score 17; DB 1; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 883 aggaagtcacacccctt 899
|||||

Db 969 AGGAGTCAACCTCCTT 953

RESULT 44

US-08-464-523B-5/C
Sequence 5, Application US/08464523B
Patent No. 5723761
GENERAL INFORMATION:
APPLICANT: Toni A. Voelker
APPLICANT: Ling Yuan
APPLICANT: Jean Kridl
APPLICANT: Deborah Hawkins
APPLICANT: Audrey Jones
TITLE OF INVENTION: Plant Acyl ACP Thioesterase
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1(a)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464,523B
FILING DATE: 02-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/13131
FILING DATE: 10-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/152,004
FILING DATE: 10-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/261,695
FILING DATE: 16-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 100-1WO
TELECOMMUNICATION INFORMATION:

TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-464-523B-5

Query Match 0.6%; Score 17; DB 1; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 883 agagatcaaccctcctt 899
|||||
DB 969 AGGAGTCAACCTCCTT 953

RESULT 45
US-08-469-203A-1/c
Sequence 1, Application US/08469203A
Patent No. 5736572
GENERAL INFORMATION:
APPLICANT: Voelker, Toni Alois
APPLICANT: Davies, Huw Maelor
TITLE OF INVENTION: Plant Medium-Chain Thioesterases
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: California
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
COMPUTER: Apple Macintosh IIfx
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,203A
CLASSIFICATION: 800
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/142,473
FILING DATE: 18-NOV-1993
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/04332
FILING DATE: 21-MAY-1992
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/824,247
FILING DATE: 22-JAN-1992
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/782,263
FILING DATE: 24-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/773,096
FILING DATE: 7-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/704,861
FILING DATE: 21-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler

REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 82-7 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 916-753-6313
TELEFAX: 916-753-1510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-469-203A-1

Query Match 0.6%; Score 17; DB 1; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 883 agagatcaaccctcctt 899
|||||
DB 969 AGGAGTCAACCTCCTT 953

RESULT 46
US-08-469-203A-1/c
Sequence 1, Application US/08469203A
Patent No. 5807893
GENERAL INFORMATION:
APPLICANT: Voelker, Toni Alois
APPLICANT: Davies, Huw Maelor
TITLE OF INVENTION: Plant Medium-Chain Thioesterases
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: California
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
COMPUTER: Apple Macintosh IIfx
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,203A
CLASSIFICATION: 800
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/142,473
FILING DATE: 18-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/04332
FILING DATE: 21-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/824,247
FILING DATE: 22-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/782,263
FILING DATE: 24-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/773,096
FILING DATE: 7-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/704,861
FILING DATE: 21-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler

REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 82-7 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 916-753-6313
TELEFAX: 916-753-1510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-469-203A-1

Query Match 0.6%; Score 17; DB 1; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 883 agagatcaaacctcctt 899
Db 969 AGGAGTCAAACCTCCTT 953

RESULT 47
US-08-470-204A-38/c
Sequence 38, Application US/08470204A
Patent No. 6028247
GENERAL INFORMATION:
APPLICANT: Voelker, Toni Alois
TITLE OF INVENTION: Plant Thioesterases
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: California
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB
COMPUTER: Apple Macintosh IIfx
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,204A
FILING DATE: 06-JUN-95
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/824,247
FILING DATE: 22-JAN-1992
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/782,263
FILING DATE: 24-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/773,096
FILING DATE: 7-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/704,861
FILING DATE: 21-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/02960
FILING DATE: 25-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/662,007
FILING DATE: 27-FEB-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/620,426
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/514,030
FILING DATE: 26-APR-1990

ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
REFERENCE/DOCKET NUMBER: CGNE 82-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 916-753-6313
TELEFAX: 916-753-1510
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-470-204A-38

Query Match 0.6%; Score 17; DB 3; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 883 agagatcaaacctcctt 899
Db 969 AGGAGTCAAACCTCCTT 953

RESULT 48
US-08-470-204A-41/c
Sequence 41, Application US/08470204A
Patent No. 6028247
GENERAL INFORMATION:
APPLICANT: Voelker, Toni Alois
TITLE OF INVENTION: Plant Thioesterases
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: California
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB
COMPUTER: Apple Macintosh IIfx
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/470,204A
FILING DATE: 06-JUN-95
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/824,247
FILING DATE: 22-JAN-1992
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/782,263
FILING DATE: 24-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/773,096
FILING DATE: 7-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/704,861
FILING DATE: 21-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/02960
FILING DATE: 25-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/662,007
FILING DATE: 27-FEB-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/620,426

FILED DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/514,030
FILING DATE: 26-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Elizabeth Lassen
REGISTRATION NUMBER: 31,845
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
REFERENCE/DOCKET NUMBER: CGNE 82-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 916-753-6313
TELEFAX: 916-753-1510
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 1561 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
US-08-470-204A-41

Query Match 0.6%; Score 17; DB 3; Length 1561;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 883 aggaagtcacacccctt 899
|||||
Db- 969 AGGAGTCAACCTCCTT 953

RESULT 49

US-08-933-750C-53
Sequence 53, Application US/08933750C

Patent No. 5933442

GENERAL INFORMATION:

APPLICANT: Lal, Preeti

APPLICANT: Hillman, Jennifer L.

APPLICANT: Bandman, Olga

APPLICANT: Shah, Purvi

APPLICANT: Au-Young, Janice

APPLICANT: Yue, Henry

APPLICANT: Guegler, Karl J.

APPLICANT: Corley, Neil C.

TITLE OF INVENTION: HUMAN REGULATORY MOLECULES

NUMBER OF SEQUENCES: 98

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/933,750C

FILING DATE: September 23, 1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0356 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 2031 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: HMC1NOT01
CLONE: 9337
US-08-933-750C-53

Query Match 0.6%; Score 17; DB 2; Length 2031;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2067 atcagacatcagagac 2083
|||||
Db 1873 ATCAGATCAGAGAC 1889

RESULT 50

US-09-234-613-53
Sequence 53, Application US/09234613

Patent No. 6132973

GENERAL INFORMATION:

APPLICANT: Lal, Preeti

APPLICANT: Hillman, Jennifer L.

APPLICANT: Bandman, Olga

APPLICANT: Shah, Purvi

APPLICANT: Au-Young, Janice

APPLICANT: Yue, Henry

APPLICANT: Guegler, Karl J.

APPLICANT: Corley, Neil C.

TITLE OF INVENTION: HUMAN REGULATORY MOLECULES

NUMBER OF SEQUENCES: 98

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/234,613

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/933,750

FILING DATE: September 23, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0356 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 53:

SEQUENCE CHARACTERISTICS:

LENGTH: 2031 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: HMC1NOT01

CLONE: 9337

US-09-234-613-53

Query Match 0.6%; Score 17; DB 3; Length 2031;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 2067 atcagacatcagaagac 2083
|||||
Db 1873 ATCAGACATCAGAGGAC 1889

Search completed: May 16, 2002, 10:48:29
Job time: 8744 sec

